

# Searches for Supersymmetry in Multi-lepton Final States with the ATLAS Detector and Related Challenges

**Katarina M. Pajchel**  
Department of Physics  
University of Oslo  
Norway



Dissertation presented for the degree of  
Philosophiae Doctor (PhD) in Physics

April 2010

© Katarina M. Pajchel, 2010

*Series of dissertations submitted to the  
Faculty of Mathematics and Natural Sciences, University of Oslo  
No. 990*

ISSN 1501-7710

All rights reserved. No part of this publication may be  
reproduced or transmitted, in any form or by any means, without permission.

Cover: Inger Sandved Anfinsen.  
Printed in Norway: AiT e-dit AS.

Produced in co-operation with Unipub.  
The thesis is produced by Unipub merely in connection with the  
thesis defence. Kindly direct all inquiries regarding the thesis to the copyright  
holder or the unit which grants the doctorate.

## Abstract

The ATLAS experiment at the Large Hadron Collider (LHC) is designed to address the open questions relating to high energy particle physics. Challenging experimental conditions require a worldwide distributed computing environment which allows the sharing of the work load and provides all collaboration member access to data and computing resources. NorduGrid is one of the three Grid flavors used by physicists for the LHC experiments. This thesis will describe the implementation of the ATLAS computing model in the NorduGrid environment. This computing infrastructure is an essential tool used for facilitating physics analysis.

One of the main goals of the ATLAS physics program is to search for Supersymmetric extensions of the Standard Model. In many cases such models predict final states with three or more leptons. This thesis investigates the potential of the ATLAS detector to discover Supersymmetry in trilepton final states. The analysis follows a jet inclusive and exclusive path aimed at two different types of trilepton events. The area of lepton isolation has been studied in more detail leading to some suggestions for optimization of the isolation requirements. Analysis strategies developed for the constrained SUSY model with gravity mediated symmetry breaking, mSUGRA, are applied on a novel set of points generated within the MSSM24 model.

The first collision data collected in December 2009 and corresponding Monte Carlo simulation has been compared in order to obtain a preliminary impression of the understanding of the detector. Outreach activities carried out during the PhD project are also reported in the thesis.





## Acknowledgements

It is perfectly allowed to joke about the length of the thesis and I have a long list of people whom I must thankfully “blame” for that. First of all, I wish to express my deepest gratitude to my supervisors Prof. Farid Ould-Saada and Prof. Lars Bugge for continuous support, both in scientific and practical matters. A special thanks to Prof. Alex Read who introduced me to the ATLAS computing jungle. The Experimental Particle Physics group at the University of Oslo has been absolutely invaluable throughout the whole project. I am mostly grateful to each and every one of you for making it a great place to work. I wish in particular to thank Bjørn H. Samset, Børge K. Gjelsten, Yuriy Pylypchenko and Samir Ferrag for close collaboration and interesting discussions. I also owe a special thanks to Jon K. Nilsen, my office-mate, whose sense of humor helped through many frustrating moments.

Parts of this work have been done in the context of the CSC analysis and I thank the members of ATLAS trilepton subgroup for a very fruitful collaboration and exchange of ideas. I also thank colleagues related to the NorduGrid collaboration for all generous help which often truly lived up to the spirit of 24/7 support.

“No man is an island”<sup>1</sup> – Through the years working with this thesis I have experienced how true this is. I wish to thank all my friends for encouragement and for being so faithful even through my most absentminded periods. My heartfelt gratitude goes to my parents and brothers who were always there challenging and encouraging me. Finally, I wish to express my deep and humble gratitude to my Dominican community at Katarinahjemmet, for their trust and generous support in so many ways, for all their sacrifices, and for bearing up with me.

---

<sup>1</sup>John Donne (1572-1631)

# Contents

|          |  |           |
|----------|--|-----------|
| <b>I</b> | <b>The Physics Challenge</b>                                 | <b>1</b>  |
| <b>1</b> | <b>General introduction</b>                                  | <b>3</b>  |
| 1.1      | Background . . . . .   | 3         |
| 1.2      | The way we work . . . . .                                    | 4         |
| 1.3      | The ingredients . . . . .                                    | 5         |
| 1.3.1    | The leptons . . . . .  | 5         |
| 1.3.2    | The quarks . . . . .   | 7         |
| 1.4      | The continuation . . . . .                                   | 10        |
| 1.5      | The experiments . . . . .                                    | 11        |
| <b>2</b> | <b>The Standard Model - a formal introduction</b>            | <b>13</b> |
| 2.1      | The fundamental SM fields . . . . .                          | 13        |
| 2.2      | The gauge principle . . . . .                                | 15        |
| 2.3      | Non-abelian gauge theory . . . . .                           | 16        |
| 2.4      | The unbearable lightness of gauge bosons . . . . .           | 18        |
| 2.4.1    | Building the full Standard Model . . . . .                   | 22        |
| 2.5      | Weak interactions and the CKM matrix . . . . .               | 26        |
| <b>3</b> | <b>Beyond the Standard Model</b>                             | <b>29</b> |
| 3.1      | SUSY chiral fields . . . . .                                 | 31        |
| 3.2      | Simple SUSY Lagrangian . . . . .                             | 31        |
| 3.3      | SUSY algebra . . . . .                                       | 32        |
| 3.4      | Towards the Wess-Zumino model . . . . .                      | 33        |
| 3.5      | SUSY as a gauge theory . . . . .                             | 34        |
| 3.5.1    | Towards the Minimal Supersymmetric Standard Model - MSSM . . | 36        |
| 3.6      | The Minimal Supersymmetric Standard Model MSSM . . . . .     | 37        |
| 3.6.1    | Spontaneous SUSY breaking . . . . .                          | 38        |
| 3.6.2    | Gauge coupling unification . . . . .                         | 39        |
| 3.7      | SUSY particle masses . . . . .                               | 41        |
| 3.8      | R-parity . . . . .   | 44        |
| 3.9      | SUSY searches, limits and mSUGRA constraints . . . . .       | 44        |
| <b>4</b> | <b>The experimental challenge</b>                            | <b>47</b> |
| 4.1      | The LHC accelerator . . . . .                                | 47        |
| 4.1.1    | The LHC magnets . . . . .                                    | 48        |
| 4.1.2    | The LHC accelerator complex . . . . .                        | 49        |
| 4.2      | ATLAS detector . . . . .                                     | 50        |
| 4.2.1    | The Inner Detector . . . . .                                 | 52        |
| 4.2.2    | Calorimetry . . . . .  | 54        |

|  |  |               |
|--|--|---------------|
| 4.2.3                                      | The magnet system . . . . .                                    | 56            |
| 4.2.4                                      | The Muon Spectrometer . . . . .                                | 56            |
| 4.2.5                                      | ATLAS forward detectors . . . . .                              | 58            |
| 4.3  | ATLAS Trigger and Data Acquisition - TDAQ . . . . .            | 58            |
| 4.4  | ID cooling . . . . .   | 59            |
| 4.5  | ATLAS commissioning and readiness . . . . .                    | 61            |
| 4.5.1                                      | Outlook . . . . .  | 62            |
| <br><b>II Computing and Data Challenge</b> |  | <br><b>63</b> |
| <b>5</b>                                   | <b>ATLAS computing</b>   | <b>65</b>     |
| 5.1  | Grid computing . . . . .                                       | 66            |
| 5.2  | ATLAS Grid computing . . . . .                                 | 67            |
| 5.2.1                                      | ATLAS data model . . . . .                                     | 67            |
| 5.2.2                                      | Data replication and distribution . . . . .                    | 69            |
| 5.2.3                                      | ATLAS Distributed Data Management . . . . .                    | 70            |
| 5.2.4                                      | ATLAS managed production . . . . .                             | 70            |
| 5.2.5                                      | Distributed Analysis . . . . .                                 | 70            |
| 5.2.6                                      | Monitoring . . . . .   | 71            |
| 5.3  | NorduGrid . . . . .  | 71            |
| 5.3.1                                      | ARC architecture . . . . .                                     | 72            |
| 5.3.2                                      | Work flow of an ARC-job . . . . .                              | 73            |
| 5.4  | NorduGrid/NDGF end-to-end Grid computing environment . . . . . | 75            |
| 5.4.1                                      | Data Managements - NDGF . . . . .                              | 75            |
| 5.4.2                                      | NorduGrid HEP Computing 2004 - 2009 . . . . .                  | 76            |
| 5.5  | NorduGrid Logger . . . . .                                     | 77            |
| 5.5.1                                      | Examples . . . . .   | 77            |
| 5.5.2                                      | Conclusions . . . . .  | 79            |
| 5.6  | RPM-distributed ATLAS software . . . . .                       | 79            |
| 5.6.1                                      | The paths of ATLAS . . . . .                                   | 80            |
| 5.6.2                                      | The ATLAS software . . . . .                                   | 80            |
| 5.6.3                                      | Installation and validation . . . . .                          | 82            |
| 5.6.4                                      | KitValidation . . . . .  | 83            |
| 5.6.5                                      | Conclusions and outlook . . . . .                              | 83            |
| 5.7  | ATLAS Managed production System . . . . .                      | 84            |
| 5.7.1                                      | Monte Carlo simulation chain . . . . .                         | 84            |
| 5.7.2                                      | ATLAS Production System . . . . .                              | 86            |
| 5.7.3                                      | Dulcinea . . . . .   | 88            |
| 5.8  | Production System operation and development . . . . .          | 94            |
| 5.8.1                                      | Storage and data management . . . . .                          | 94            |
| 5.8.2                                      | Job Error analysis . . . . .                                   | 95            |
| 5.8.3                                      | Operation issues and tools . . . . .                           | 97            |
| 5.8.4                                      | Resource and performance statistics . . . . .                  | 100           |
| 5.8.5                                      | Outlook . . . . .  | 104           |
| 5.9  | ATLAS distributed analysis - Ganga . . . . .                   | 104           |
| 5.9.1                                      | Ganga architecture . . . . .                                   | 105           |
| 5.9.2                                      | GangaNG - the NorduGrid backend . . . . .                      | 106           |

|            |  |            |
|------------|--|------------|
| 5.9.3      | Installation and configuration . . . . .                                     | 109        |
| 5.9.4      | NG backend object . . . . .  | 110        |
| 5.9.5      | Conclusion . . . . .   | 112        |
| <b>III</b> | <b>The Analysis Challenge</b>  | <b>113</b> |
| <b>6</b>   | <b>Phenomenology - SUSY and the Standard Model</b>                           | <b>115</b> |
| 6.1        | Sparticle production . . . . .   | 115        |
| 6.1.1      | Gaugino production . . . . .   | 116        |
| 6.1.2      | Squark and gluino production . . . . .                                       | 117        |
| 6.1.3      | Gaugino decays . . . . .   | 119        |
| 6.1.4      | Gluino and squark decays . . . . .   | 121        |
| 6.1.5      | Slepton decays . . . . .   | 121        |
| 6.2        | Constrained MSSM – mSUGRA . . . . .  | 122        |
| 6.3        | Signal and background in mSUGRA searches . . . . .                           | 123        |
| 6.3.1      | Co-annihilation Region (SU1) . . . . .                                       | 124        |
| 6.3.2      | The Focus Point Region (SU2) . . . . .                                       | 124        |
| 6.3.3      | The Bulk Point Region (SU3) . . . . .  | 125        |
| 6.3.4      | The Low Mass Point Region (SU4) . . . . .                                    | 125        |
| 6.3.5      | Standard Model Background . . . . .  | 126        |
| 6.4        | Monte Carlo Simulation . . . . .   | 128        |
| 6.5        | Analysis data preparation . . . . .  | 130        |
| <b>7</b>   | <b>Physics Object reconstruction and particle identification performance</b> | <b>131</b> |
| 7.1        | Jets and missing energy . . . . .  | 131        |
| 7.2        | Muon isolation and identification study . . . . .                            | 132        |
| 7.2.1      | Muon MC-truth information . . . . .  | 133        |
| 7.2.2      | Muon isolation study . . . . .   | 135        |
| 7.2.3      | Muon-jet separation . . . . .  | 138        |
| 7.3        | Electron isolation and identification study . . . . .                        | 142        |
| 7.3.1      | Electron MC-truth information . . . . .                                      | 144        |
| 7.3.2      | Electron isolation study . . . . .   | 145        |
| 7.3.3      | Electron-jet separation . . . . .  | 148        |
| 7.4        | Isolation optimization for trilepton signal . . . . .                        | 150        |
| 7.5        | Single lepton selection efficiency and fake rate . . . . .                   | 152        |
| 7.5.1      | Performance of the different isolation requirements . . . . .                | 155        |
| <b>8</b>   | <b>SUSY searches, trilepton analysis at 14 TeV</b>                           | <b>159</b> |
| 8.1        | Object definition . . . . .  | 159        |
| 8.1.1      | Significance measures . . . . .  | 161        |
| 8.2        | Jet inclusive analysis . . . . .   | 162        |
| 8.2.1      | “3-leptons+jet” . . . . .  | 163        |
| 8.2.2      | “3-leptons+ $\cancel{E}_T$ ” . . . . .                                       | 165        |
| 8.2.3      | Jet inclusive analysis - summary, conclusions . . . . .                      | 167        |
| 8.3        | Jet exclusive analysis . . . . .   | 168        |
| 8.3.1      | Jet exclusive analysis - summary, conclusions . . . . .                      | 171        |
| 8.3.2      | Trigger efficiency for 3-lepton events . . . . .                             | 172        |

|           |  |            |
|-----------|--|------------|
| 8.4       | Preliminary Conclusions and Outlook . . . . .                | 172        |
| <b>9</b>  | <b>Trilepton SUSY searches with early data at 10 TeV</b>     | <b>175</b> |
| 9.1       | mSUGRA benchmark points at 10 TeV . . . . .                  | 175        |
| 9.2       | Standard Model background samples at 10 TeV . . . . .        | 176        |
| 9.3       | Object definition . . . . .                                  | 177        |
| 9.3.1     | Lepton jet separation . . . . .                              | 178        |
| 9.3.2     | Anti- $K_T$ Jets . . . . .                                   | 180        |
| 9.4       | Significance . . . . .                                       | 183        |
| 9.5       | Jet inclusive trilepton searches . . . . .                   | 184        |
| 9.5.1     | Variation of the jet inclusive analysis . . . . .            | 188        |
| 9.6       | Conclusions . . . . .  | 191        |
| <b>10</b> | <b>MSSM24 multilepton analysis with early data at 10 TeV</b> | <b>193</b> |
| 10.1      | Gaungino production and trilepton signal in MSSM24 . . . . . | 194        |
| 10.1.1    | Jet exclusive analysis . . . . .                             | 196        |
| 10.1.2    | Alternative track based isolation . . . . .                  | 199        |
| 10.2      | Gauginoshake . . . . .                                       | 200        |
| 10.2.1    | Jet inclusive analysis . . . . .                             | 201        |
| 10.2.2    | 4-lepton signal . . . . .                                    | 202        |
| 10.3      | Conclusions . . . . .  | 204        |
| <b>11</b> | <b>First collision data</b>                                  | <b>207</b> |
| 11.1      | Event selection . . . . .                                    | 207        |
| 11.2      | Jets and missing $E_T$ . . . . .                             | 209        |
| 11.3      | Tracks . . . . .   | 210        |
| 11.4      | Muons . . . . .  | 210        |
| 11.4.1    | Muon isolation . . . . .                                     | 214        |
| 11.5      | Electrons . . . . .  | 215        |
| 11.5.1    | Electron isolation . . . . .                                 | 219        |
| 11.5.2    | Pions as electrons . . . . .                                 | 224        |
| 11.6      | Conclusions and outlook . . . . .                            | 226        |
| <b>12</b> | <b>Systematic effects</b>                                    | <b>227</b> |
| 12.1      | Systematic effects - general comments . . . . .              | 227        |
| 12.1.1    | MC related effects . . . . .                                 | 227        |
| 12.1.2    | Significance measures . . . . .                              | 228        |
| 12.1.3    | Trigger . . . . .  | 229        |
| 12.2      | 10 TeV jet inclusive analysis . . . . .                      | 229        |
| 12.2.1    | Cut optimization . . . . .                                   | 229        |
| 12.2.2    | Variation of cuts around the working point . . . . .         | 230        |
| 12.2.3    | Systematic effects: 10 TeV jet inclusive analysis . . . . .  | 234        |
| 12.3      | 14 TeV analysis . . . . .                                    | 237        |
| 12.3.1    | Overall summary jet inclusive analysis at 14 TeV . . . . .   | 239        |
| 12.3.2    | 14 TeV jet exclusive analysis . . . . .                      | 239        |
| 12.4      | Conclusions . . . . .  | 241        |

---

|   |            |
|---|------------|
| <b>IV The Outreach Challenge</b>  | <b>247</b> |
| <b>13 Outreach activities</b>   | <b>249</b> |
| 13.1 Masterclass 2005-2008 . . . . .  | 249        |
| 13.2 Department booth at National Science Fair 2008 “What are you made of?” | 252        |
| 13.3 Radio lecture: “There has to be something more” . . . . .              | 254        |
| 13.4 Outlook . . . . .  | 254        |
| <b>A ATLAS computing</b>  | <b>257</b> |
| A.1 Environment variables for ATLAS building . . . . .                      | 257        |
| A.2 External packages . . . . .   | 258        |
| A.3 Installation of ATLAS external packages . . . . .                       | 258        |
| A.4 KitValidation . . . . .   | 259        |
| A.5 ATLAS production database ProdDB . . . . .                              | 262        |
| <b>B SUSY branching fractions</b>   | <b>265</b> |
| <b>C List of 10 TeV MC samples</b>  | <b>267</b> |
| <b>D List of outreach contributions</b>                                     | <b>271</b> |
| <b>E List of papers</b>   | <b>273</b> |

# Part I

## The Physics Challenge





# Chapter 1

## General introduction

The general outline of this thesis is to first set the scenes, present the questions we are facing in particle physics and the search for answers by describing the experiment itself, the apparatus (Chapter 4), and the computing tools we use (Part II). This is followed by a presentation of some possible search methods and results based on Monte Carlo simulations (Part III) which indicate how physicists will interpret the data. The study is based on collisions which produce three or more leptons in the final state. Observing a larger number of such events than we expect according to the current knowledge, could be an indication of processes of new and broader theories. Trilepton events are one of many possible signatures of the extended Supersymmetric theories of particle physics. Such events may be rare, but they are clear and there are few other known Standard Model processes that give a similar signature in the detector, making them a strong hint of something new. Some strategies were found to possibly be fruitful already during the analysis of early data collections while others require larger data samples that will only be available after several years. ATLAS collected its first collision data in November and December 2009. Chapter 11 shows an early study of the 900 GeV collision data.

The last part of the thesis (Part IV) looks out towards a broader context and reports on the outreach activities carried out during the PhD project.

### 1.1 Background

The material presented in a theory introduction for a particle physics thesis is a summary of the writer's knowledge gleaned from courses and literature. The challenge is of course to choose a good selection of the material and present it in a clear way, possibly with an original twist. In the present case, part of the project has been general public outreach. It is therefore hoped that the reader accepts this choice of specific flavor for this introduction with its very basic character aimed at an audience without a particle physics background. It is based on the writer's experience in the field of outreach gained during her PhD project and reflects aspects of this work. Similar introductions have been presented in a radio lecture and several public lectures [1–3].

In writing for a general public audience the challenge is to explain how physicists work, the very process of scientific endeavor, physicists' motivation, questions currently being asked and struggles in finding answers. While physicists share their own excitement and fascination, they also have to justify and explain why this effort is interesting, valuable and important, in relation to a broader scientific, cultural and social perspective.

The outline of the general introduction is as follows: Section 1.2 describes the process of particle physics research with its interaction between theory and experiment, Section 1.3 introduces the various building blocks and forces found in Nature following a historical path. In Section 1.4 our attention is turned toward the starting experiments at the Large Hadron Collider (LHC) and the searches for new phenomena. Finally, Section 1.5 presents the basic experimental methods of high energy physics.

### 1.2 The way we work

Particle physics as the quest for fundamental building blocks of Nature finds its roots in the ancient Greece where the philosopher Democritos (460-370 B.C.) claimed that all things were built up of indivisible particles, *atomos*. The objects studied in particle physics today are much smaller than the building blocks that are currently called atoms. However, the idea and the questions are the same as in the antique: What are the fundamental constituents and laws of matter?

In order to better understand how we explain the various processes, why we claim the things we claim and how we can convince even the funding agencies to actually invest in our experiments, let us look for a moment at how particle physicists work. Roughly speaking, they fall in two categories: theoreticians and experimentalists.

Theoretical physicists use fascinating mathematics as the language to describe all particles and forces within a framework which is called quantum field theory. One of the things that has really puzzled several of the greatest physicists is that Nature actually can be understood, even in great depth and detail, and it can be described by mathematics. Mathematics is after all in its very foundation built on logic and created by human beings which could in a certain sense allow us to call it a humanistic discipline.

So the theoreticians try to build mathematical models that describe the particles and forces in a correct way. The formulas are like the DNA code of the theory which contain information about all particles and what processes they are involved in, how they are created and how they decay. There is room for much creativity, but of course, the theories must be confronted with experiments in order to decide what is right and what is wrong. At the same time, the theories must also respect the laws of mathematics and the their mathematical structures. These are actually some of the most fruitful guidelines in the development of new ideas. During the development of particle physics in the 20th century we have seen many examples of theoreticians who, trusting that Nature is mathematical, could say: “given this is correct” and “if this equation is to be solved” or “in order to complete this picture”, then we expect to observe certain new particles or processes.

That is where the experimentalists enter the scene and try to construct experiments and analyses that as unambiguously as possible can decide what is the correct theory. Often there are many models on the market and one has also to bear in mind that the picture is changing depending on the energy conditions (scale) of the experiment. Time and again we see that theories or descriptions of phenomena are not wrong, but rather we learn that they have a limited validity and thus we have to look for a broader, more general picture. Such theories with a limited validity range are often called *effective theories*.

But of course, despite the prophetic voices of the theoreticians, Nature is always capable of surprising with things nobody had thought of. There is always a two way communication. The experimental results are a continuous input in the theory calculations and models.

It is exactly this fascinating dialog between the mathematical theories and experiments that has unfolded before us the overwhelming picture of some of the fundamental constituents of Nature. And yet we are far from done. In many cases we ask why something can happen at all. An answer of the type “quantity X is negative” may be perfectly correct, nevertheless, it does not make us much wiser. Next question would naturally be: Why is it negative? Therefore we continue to seek for ever deeper and more general insight.

## 1.3 The ingredients

All matter can at the most fundamental level be described as built up of indivisible particles. How these are created and decay, how they interact with each other, how they come together and create new particles is governed by fundamental forces which will be described in a moment. Not all types of particles can “feel” all of them. Which category a particle belongs to is very closely related to which forces it “feels” and this fact defines the classes of particles. Just by knowing which type a particle belongs to, information is gleaned about which processes it can get involved in.

The notion of *force* is familiar from everyday life. But in the quantum mechanical world of particle physics it is more correct to speak about *interactions*, and these are mediated by special particles which are called *force carriers*. One could imagine an interaction between two particles as two persons on skates that throw a ball to each other. The person who throws the ball will experience being pushed backwards. The person who catches it will have the same experience. One can say that the ball has transferred some energy between these two.

Energy is another key notion and is found in nature in many different forms. It is the foundation of all there is, the actual raw material. Einstein’s famous formula

$$E = mc^2$$

shows the relation between energy  $E$  and mass  $m$ . If one wants to create a particle with mass  $m$  one needs an energy which is equivalent to the mass multiplied with the speed of light  $c$  squared. As the speed of light is very high,  $3 \times 10^8 \text{m/s}$ , one needs a lot of energy to create particles with substantial mass. On the other hand, mass can also be transformed into energy and in effect release large amounts of energy. Some examples: the nuclear reactions in the Sun as well as in nuclear plants, provide energy according to this relation.

### 1.3.1 The leptons

After this introduction of the fundamental concepts, let us see how the strange world of particles unfolded during the last 100 years.

The first glimpse of it appeared in 1897 when Thomson discovered the electron. During the first 35 years of the 20th century, one had identified all particles that build up the ordinary matter as seen around us. These are the lightest particles in their respective classes. As they are light, they are also stable which means they do not disintegrate, or in other words decay to other particles. For the sake of completeness, it is the electron, and the proton and neutron. The two last-named build up the nucleus of the atom. This was a great success, but of course there was the ever present question: Are these particles fundamental or are they made up of even smaller particles?

The early discoveries were done through observations of radioactive sources, which provided particles with relatively low energy, or cosmic radiation. The Space is a constant source of highly energetic particles which collide with the atoms of the atmosphere and shower the earth with different sorts of particles. It was in the cosmic ray studies that one in 1937 observed a new particle that seemed to be just like the electron only 200 times heavier and apparently did not have any purpose. The surprise was very well formulated by Isodor Isaak Rabi who upon hearing the news asked: “Who ordered that?”.

During these years we have two interesting examples of how theory and observation mutually interacted. In 1927-28 Paul Dirac laid the foundation of the quantum theory of electrodynamics, also called QED. The theory described interactions between particles carrying electrical charge. He also formulated the relativistic equation describing the electrons.

In his own view, it was a beautiful and mathematically compelling equation, and it had a great predictive power. Nevertheless, it provided a solution which despite its mathematical correctness, seemed to be unphysical. In physics there are several examples of such redundant solution, but in this case Dirac claimed that it had a meaning and he interpreted it as an electron with an opposite electric charge, a so-called positron. Four years later in 1932 Carl D. Anderson made observations of a particle that could only be interpreted as an electron with positive charge. This was at the same time a first evidence of anti-matter.

Another example of theory preceding observation is found in the history of the mysterious neutrino. Wolfgang Pauli proposed in 1930 that the missing energy measured in the decay of neutrons into a proton and electron, a so called  $\beta$ -decay, could be explained by an unobserved new particle. Since it would have to be electrically neutral and light it was named by Enrico Fermi *neutrino*, the *little neutral one*, however, in other respects it is an electrically neutral friend of the electron. Fermi published also in 1934 a more complete theory of beta-decay which included the neutrino. The idea was considered so speculative that the scientific journal Nature refused publishing it.

Around 1950 one had discovered several processes which were assumed to be of the same nature as the  $\beta$ -decay and also described by Fermi's theory which became known as the *weak interaction*. The existence of the neutrino was experimentally confirmed first in 1953 by Frederick Reines and his colleague Clyde L. Cowan, Jr. The problem with neutrinos is that they really live up to their name. They can penetrate matter barely noticing it exactly because they interact only through the weak interaction. Actually a very large number of neutrinos from space is passing undisturbed through us and through the Earth.

But sometimes Nature takes the lead. In 1960 one discovered at the Brookhaven laboratory that the muon, the “Who-ordered-that” heavy sibling of the electron also had an illusive neutrino partner. The confusion was ever growing when the SPEAR experiment at the Stanford Linear Accelerator (SLAC) in 1975 reported that they had discovered an even heavier “electron”. It was a purely experimental adventure where one did not really have any theoretical reason for expecting a new particle. One wanted to explore the unknown, and Nature surprised.

The new heavy particle was called *tau*,  $\tau$ . It is 17 times as heavy as the muon and since it appeared to be an electron-like particle, one assumed that also the tau had its own neutrino. The argument was that the picture would be incomplete without it. It was finally completed in 2000 when the existence of the tau neutrino was confirmed at

Fermilab, the high energy physics laboratory outside Chicago. There are very strong evidence that there are not more than three families, each with an electron-like particle and its neutrino partner. As a group they are all called leptons, “the light ones” in Greek. Which interactions do they get involved in? The electron-like particles, the electron, muon and tau, are charged and interact via the electromagnetic force which is mediated by the massless force carrier particle *photon*. The photon is actually nothing else than the very familiar light particle, the only particle we can actually “see”. The electrically neutral neutrino is excluded from that game, but the weak interaction involves all leptons. This force is mediated by particles which contrary to the photon are actually rather heavy and this force has therefore a very short range. These are called the  $W^\pm$  and  $Z^0$  particles and were finally discovered in 1983 at CERN, the European high energy physics laboratory outside Geneva, long after they had become part of the theory. Although this force is very weak compared to the electromagnetic interaction, we would not exist without it. We depend on it indirectly as it governs various nuclear processes, for example those in the Sun.

### 1.3.2 The quarks

This was so far the story of the leptons. What happened with the protons and neutrons? Protons are positively charged so there is a repulsive force between them. Nevertheless, together with the neutrons they build up the atomic nucleus and actually stick quite tightly together. The force that is responsible for keeping the nucleus together must therefore be much stronger than electromagnetism and for this reason it was called the *strong force*. In the fifties and sixties one managed to collide particles at ever growing energies and a whole spectrum of new strongly interacting particles were discovered. Physicists were overwhelmed. J. Robert Oppenheimer suggested that for once one should give the Nobel prize to a physicist who *did not* discover a new particle, while Enrico Fermi claimed “Young man, if I could remember the names of these particles, I would have been a botanist.”.

This particle zoo opposed the intuition of physicists which always goes in the direction of simple principles which involve only few fundamental factors. In 1964 Murray Gell-Mann and George Zweig proposed independently the so called quark model. It offered a classification scheme built on the assumption that these strongly interacting particles were built up of even smaller constituents: *quarks*. Such particles built up of quarks are denoted by the common name *hadrons*, yet another word from Greek, meaning the “thick” or “heavy”. The quarks were assumed to have fractional electrical charge ( $\pm 2/3$  or  $\mp 1/3$ ) and other quantum mechanical features which defined how they coupled in pairs or triplets. If a particle is made of two quarks, the combination must be such that the charges add up to  $+1$ ,  $-1$  or  $0$  in which case it is electrically neutral. Such particles are called mesons. For example, the  $\pi$ -meson, is built up of the two lightest quarks: up and down. A positively charged pion  $\pi^+$  is made of an up quark ( $u$ , charge  $+2/3$ ) and anti down (denoted by  $\bar{d}$  and with charge  $+1/3$ ), so  $u\bar{d}$ . To get the negative pion one simply flips the charge of the quarks and a  $\pi^-$  is made of  $\bar{u}d$ . In addition there is also a neutral pion  $\pi^0$  made of a mixture of  $u\bar{u}$  and  $d\bar{d}$ . In the fifties one started observing a new particle with a mass almost four times the pion mass, relatively long lifetime and it decayed into pions. Abraham Pais postulated a new quantum number called *strangeness* related to the heavier *strange quark*  $s$  (charge  $-1/3$ ). This quark could explain the appearance of a new family of mesons called  $K$ -mesons.

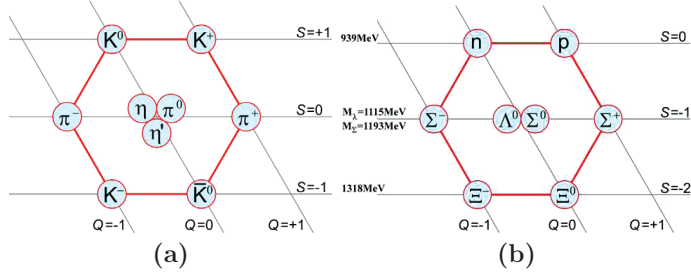


Figure 1.1: (a): Nonet of the lightest mesons . (b): Octet of the lightest baryons.

Quarks can also make up triplets which are called *baryons*. Most familiar are the proton made of  $uud$  (charge 1) and the neutron  $udd$  (charge 0). Take now the proton and change the  $d$  with an  $s$  quark, and one gets the slightly heavier  $\Sigma^+$  particle (charge 1).

With the three quarks one could group the observed composite particles in symmetry patterns, in groups of 8 or 9. Figure 1.1 shows on the left the nonet containing the lightest mesons, among them the  $\pi$  and  $K$ -mesons. On the right there is the octet of the lightest baryons with the familiar proton and neutron in the upper row.

In the beginning the schemes had some holes, but the particles predicted by the quark model were actually found, even with the anticipated mass and charge. This is again a good example of how theory could point ahead and the discoveries were considered as great support of the quark model.

One of the most compelling evidence of the quarks was the discovery of the so called  $J/\psi$  particle in 1974. By this time one knew of three quarks, the up ( $u$ ), down ( $d$ ) and strange ( $s$ ). The  $J/\psi$  was much heavier and had a longer lifetime than the other known unstable hadrons. The hypothesis was that it had to contain a new type of quarks, a so called charm quark ( $c$ ) and that it was a bound state of a  $c$  and anti- $c$ . If this was correct, one could expect a whole new family of particles where the charm formed particles with the three already known quarks ( $u, d$  and  $s$ ). If these were found, it would be a great confirmation of the model. As the experiments developed and could study collisions at ever higher energies one could once again build up a beautiful pattern of new and heavier hadrons. After the discovery of the  $c$  quark one could set up two complete families of quarks, two in each, just like for the leptons.

The confirmation of the existence of the quarks was strong, despite the fact that no one had observed them directly. Quarks are said to be the most social particles as they are always found in groups of two or three and never alone. This is explained by the nature of the strong force that keeps them together. It is mediated by the massless *gluon* particle which acts almost like a rubber band. The farther one pulls two quarks apart from each other, the stronger the force between them. If one “pulls” really hard, instead of two free quarks, the energy which is put into this process will turn into new quarks that gather into new bound states. This is actually the production mechanism responsible for the hadron-zoo. However, this productivity had to follow some unbreakable quantum mechanical rules. Quarks with exactly the same properties can not form a bound state. Some quantum mechanical property has to make them unique in the group. When the experiments on the contrary discovered combinations of quarks which seemed to be identical, one had to introduce a new quantum number which could distinguish them. It was called “color”: red, green and blue. This color charge which is assigned to quarks

| Particles                         | Weak | EM  | Strong | Gravity |
|-----------------------------------|------|-----|--------|---------|
| Neutral leptons,<br>neutrinos     | Yes  | No  | No     | “Yes”*  |
| Charged leptons<br>$e, \mu, \tau$ | Yes  | Yes | No     | “Yes”*  |
| Quarks                            | Yes  | Yes | Yes    | “Yes”*  |

Table 1.1: Particles and their interactions. EM stands for Electromagnetism. (\*) For relatively tiny masses gravity is negligible.

and gluons is an analogy to the color theory in optics. Bound states of quarks are always color neutral, so-called white. Like in electrodynamics where the charges define repulsion or attraction, the color charge is the key to the understanding of gluon mediated strong interactions. Therefore, the theory describing the strong interactions involving quarks and gluons is called *quantum chromodynamics* (QCD) using the Greek word for color “chroma”.

The discoveries in the hadron world did not end with the  $c$  quark. In 1977 Fermilab studied proton collisions which produced muon pairs that showed clear evidence of a new particle with a well defined mass. These were suggested to involve an entirely new quark which was called bottom, or  $b$ -quark. Like in the previous cases, one could study a whole new rich spectrum of new particles in a higher mass range. Although it could fit into the quark model, one had no explanation why there should be a third family or if there were more than three families. But the model was already convincing enough to predict that the third family also consisted of two quarks where the  $b$ -quark was accompanied by a heavier quark called “top” ( $t$ ). And indeed, it is heavier. It weights more than a silver atom and it was not until 1995 that one had the technology to collide particles at energies high enough to afford to make a pair of such heavy beasts. This discovery was also done at Fermilab.

In the nineties (1989-2000) the Large Electron Positron (LEP) accelerator at CERN made a number of precise measurements which involved all known particles, except the top quark. One of the main objects of study was the  $Z^0$ -particle, one of the weak force carriers. Observations of how it decayed provided very strong evidence that there are only three light families of leptons and quarks. The discoveries of the particles in the third families of leptons and quarks, together with the studies of force mediating particles, so-called *bosons*, completed the picture which we today call the Standard Model (SM) of matter and forces.

To summarize the journey so far. The Standard Model of particle physics includes three forces: electromagnetism, the weak and the strong, each mediated by a force carrier particle - photon for electromagnetism,  $W^\pm$  and  $Z^0$  for the weak and gluon for the strong force. The matter particles are classified in two main groups: leptons and quarks, each consisting of three families with higher masses for each. Figure 1.2 shows a schematic overview of the building blocks of the Standard Model, while Table 1.1 summarizes which interaction the different particles can get involved in.

The 12 matter particles have also their respective anti-particles. In Nature quarks are only found in composite particles consisting of two or three quarks and form a rich flora of so-called hadrons.



| Three Generations<br>of Matter (Fermions) |  |  |  |  |
|---|--|--|--|--|
|   | I  | II   | III  |  |
| mass→                                     | 2.4 MeV  | 1.27 GeV   | 171.2 GeV  | 0  |
| charge→                                   | $\frac{2}{3}$  | $\frac{2}{3}$  | $\frac{2}{3}$  | 0  |
| spin→                                     | $\frac{1}{2}$  | $\frac{1}{2}$  | $\frac{1}{2}$  | 1  |
| name→                                     | u<br>up  | c<br>charm   | t<br>top   | $\gamma$<br>photon                             |
| Quarks                                    | 4.8 MeV<br>$-\frac{1}{3}$<br>$\frac{1}{2}$<br>d<br>down          | 104 MeV<br>$-\frac{1}{3}$<br>$\frac{1}{2}$<br>s<br>strange       | 4.2 GeV<br>$-\frac{1}{3}$<br>$\frac{1}{2}$<br>b<br>bottom        | 0<br>0<br>1<br>g<br>gluon                      |
|   | <2.2 eV<br>0<br>$\frac{1}{2}$<br>$\nu_e$<br>electron<br>neutrino | <0.17 MeV<br>0<br>$\frac{1}{2}$<br>$\nu_\mu$<br>muon<br>neutrino | <15.5 MeV<br>0<br>$\frac{1}{2}$<br>$\nu_\tau$<br>tau<br>neutrino | 91.2 GeV<br>0<br>1<br>Z<br>weak<br>force       |
|   | 0.511 MeV<br>-1<br>$\frac{1}{2}$<br>e<br>electron                | 105.7 MeV<br>-1<br>$\frac{1}{2}$<br>$\mu$<br>muon                | 1.777 GeV<br>-1<br>$\frac{1}{2}$<br>$\tau$<br>tau                | 80.4 GeV<br>$\pm 1$<br>1<br>W<br>weak<br>force |
| Leptons                                   |  |  |  | Bosons (Forces)                                |

Figure 1.2: The Standard Model of particle physics.

## 1.4 The continuation

All experimental results so far show an astonishing agreement with the Standard Model. Nevertheless there are some open “alarming” questions. The Standard Model does not include gravity. Of course, all things that have mass, including the particles, feel gravity, however at the microscopic scale it is negligible compared to the other forces. To this day it has not been possible to formulate a mathematical theory of gravity along the same pattern of quantum fields as for the other interactions.

Another long standing challenge is the question of the origin of mass. The simplest formulation of the equations describing the Standard Model contains massless particles and every straight forward introduction of mass would ruin the theory. The mass has to be introduced in a clever way and such a mechanism was developed in 1964 by a number of physicists. Today it has become known as the Higgs mechanism. If this is the correct solution, one should be able to observe in experiments the so-called Higgs particle. The LEP experiments did not find it, but set a minimum limit on the mass of the Higgs, while the Fermilab experiments have recently excluded a higher mass interval. As the search is closing up on the Higgs, one awaits with great expectations the start-up of the Large Hadron Collider (LHC) at CERN. LHC was in fact designed to cover the whole energy range where the Higgs particle should be if the mechanism it indicates is the correct answer.

One of the most important paths of current studies is to look for extensions of the Standard Model. Despite of the great success of the model, there is a number of observations and features in the theory that strongly indicates that the simple picture of Figure 1.2 is not the whole picture. One of the most compelling theories on the market is Supersymmetry (SUSY) whose basic idea is to add a so-called supersymmetric partner to every known Standard Model particle. These partners are equal in all respects, except for one quantum mechanical property called spin and of course mass. Since we have not found them yet, it is assumed that they are heavier than the familiar Standard Model



particles. In addition to fixing several serious problems in the Standard Model, it also offers a candidate for the dark matter. This matter whose nature is as of today unknown, is assumed to make up as much as 23% of the total mass in the universe. It is called dark because it is not visible in any other way than through gravitational effects. With the present list of particles physicists can only explain the visible part of the mass of the universe, meaning basically the stars, which account for only 4% of the mass of the universe. This leaves us with the question how to explain the remaining 73%. Currently it is assigned to the elusive dark energy.

In addition there are many other exciting ideas that will be up for testing at the LHC. More fundamental theoretical frameworks suggest that space itself may have extra dimensions in addition to the four familiar ones. This would again have an impact on gravity and under special conditions would allow us to observe gravitational effects like gravitons – gravity mediating particles – or even mini black holes.

These and many other questions are on the to-do list of the physicists in the coming years. The situation at the moment is similar to what we have experienced several times earlier. There is a very convincing and beautiful theoretical framework and an impressive list of measurements which confirm the theory or contribute to more and more precise estimates of the still unknown factors. All this accumulated knowledge, both theoretical and experimental, points clearly beyond the Standard Model towards a larger and more general picture. Another fascinating feature is that the more general theories we deal with and the higher energies we study, the more important becomes the link to astrophysics, cosmology and the early evolution of the universe.

It is this insight and trust in the fruitful interplay of theory and experiment that has been a driving force behind the enormous LHC-project.

Hopefully this brief historical introduction of the Standard Model has shown how physicists work. Learning from the past, even though most of the new physics we plan to search for at the LHC are hypotheses, we strongly believe that LHC will reveal *something*, although not necessarily exactly what we prepare for, but definitely *something*.

## 1.5 The experiments

When we speak of particles with high energy, we think of particles at very high speed, but also of particles which in themselves carry a lot of energy due to their high mass. Creating heavy particles is “expensive” with respect to energy. As we are already familiar with all the lighter particles and want to create new and heavier ones, we need to somehow get hold of the energy that is required. This is done by accelerating particles to a very high speed, close to the speed of light, using an *accelerator* machine. The ones used in the most recent experiments at Fermilab and at CERN are shaped as large circles. At the LHC we send packages or so-called bunches of protons into the ring and accelerate them using electromagnetic field. The charged protons can be deflected by magnetic fields. The accelerator ring is therefore mainly built up of 15 m long magnets which basically look like thick pipes. In the very center there are two narrow beam pipes through which the protons are passing (see Figure 4.1). The magnetic fields are used to guide the protons around the 27 km long circle.

Once the protons revolve at tremendous speed in both directions, in order to “release” the energy and create new particles, they are at certain points lead to collide. In such collisions, the accumulated energy of the fast particles is transformed in various interac-

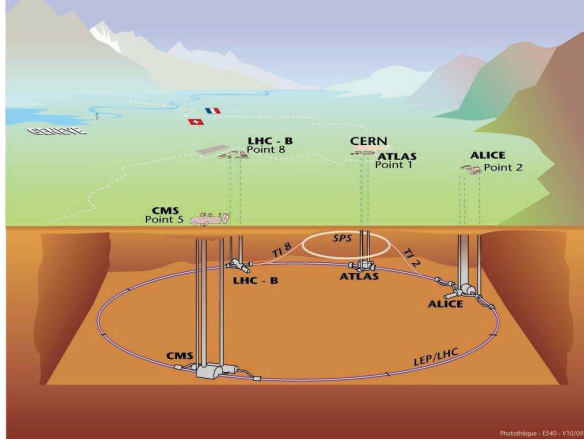


Figure 1.3: Overview of the LHC accelerator at the border between Switzerland and France approximately 100 m under ground. ATLAS is situated at Point 1.

tions into new particles. In order to observe them, the collision points are at the center of barrel shaped detectors. Figure 4.3 shows the ATLAS detector which surrounds one of the four collision points.

The rarely produced heavy particles have usually a very short lifetime and decay into other lighter particles which fly out in all directions into the detector which registers them as they pass by. A detector is like a large digital camera with millions of sensitive “pixels” throughout the whole volume. The information about the particles inside the detector is read out and tracks and energy deposits are reconstructed. There is no way to directly see what happened at the collision point, so all we can do is to use the particles registered in the detector to reconstruct the short-lived particles that were created in the collision and try to guess which type of process it was. A more detailed description of the LHC and in particular the ATLAS detector can be found in Chapter 4.

Like the other detectors at the LHC, ATLAS is designed to collect as much as possible and as precise as possible information about the particles that pass through it. This leads to enormous amounts of data which needs to be processed, stored and distributed to the physicists. In order to handle and analyze the data we need a global computing facility, a so-called Data Grid, where a number of computing centers around the world share the work load. This system will be described in more detail in Part II.

# Chapter 2

## The Standard Model - a formal introduction

Mathematically the Standard Model is formulated as a Quantum Field Theory (QFT) where interactions are introduced by the gauge invariance principle. QFT is a consistent synthesis of quantum mechanics and special relativity applied to point particles. Quantum mechanical effects become important only at a certain scale. In this theory the Planck's constant

$$\hbar = h/2\pi = 1.054\,572 \times 10^{-34} \text{ J s} = 6.582\,119 \times 10^{-22} \text{ MeV s} \quad (2.1)$$

is a fundamental quantity and it sets the maximum limit on the action  $S = \int L dt$ , where  $L$  is the Lagrangian describing the system. When action becomes large,  $S \gg \hbar$ , quantum effects are no longer important and one enters the classical domain. In special relativity, on the other hand, the speed of light  $c = 299,792,458 \text{ m/s}$  is the key quantity. Relativistic effects become significant only as we approach this ultimate speed limit.

This presentation is based on references [4–9].

### 2.1 The fundamental SM fields

The ingredients of the theory are particles characterized by properties like mass, different types of charges and spins, statistical properties and chirality. The last property is related to the fact that certain phenomena are not invariant under parity transformations which can be understood as mirroring. Chirality is strongly related to the helicity of particles which is defined as the projection of the spin on the momentum vector. If these point in the same direction the helicity is positive and the particle is called right-handed. While if the spin and momentum point in opposite directions, the helicity is negative and the particle called left-handed.

Bosons which have zero spin and follow the Bose-Einstein statistics are described by scalar fields  $\phi(x)$ , solutions of the Klein-Gordon equation

$$(\partial^\mu \partial_\mu + m^2)\phi(x) = 0 \quad (2.2)$$

and can be written as a superposition of plane waves. After quantization, the operator coefficients of the momentum-space Fourier expansion of  $\phi(x)$  are interpreted as creation and annihilation operators and obey commutator algebra. This fact is directly linked to the Bose-Einstein statistical behaviour of the scalars. The commutators play again an

important role in the definition of the propagators of scalar particles which will enter the calculation of physical quantities like cross section.

Although we started from the Klein-Gordon equation, it is the Lagrangian density that is more important in the QFT Lagrangian formalism. The Lagrangian appropriate to the scalar field is given by

$$\mathcal{L}_{\text{K-G}} = \frac{1}{2}(\partial_\mu \phi(x) \partial^\mu \phi(x) - m^2 \phi(x)^2) = 0. \quad (2.3)$$

Scalar fields may also be complex and described by a Lagrangian that contains both the field and its complex conjugate which accounts for the antiparticle. The equation of motion, like in this case the Klein-Gordon equation (Eq. 2.2), can be obtained from the Lagrange density (Eq. 2.3) via the Euler-Lagrange equation

$$\partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0 \quad (2.4)$$

derived from the principle of least action.

The massive spin-1/2 particles, and in general all half-integer spin particles, are described by the Dirac equation

$$(i\gamma^\mu \partial_\mu - m)\psi(x) = 0. \quad (2.5)$$

The field  $\psi(x)$  satisfying this equation is a 4-component spinor field accounting for the two spin states of a particle and corresponding anti-particle. This equation has both a positive and negative energy solution. Historically the negative solution was thought of as non-physical. However, Dirac gave it a physical interpretation where the vacuum was considered as many-body quantum state, a “sea” where all negative-energy eigenstates were filled. A possible “hole” in that sea could be interpreted as an electron with a positive charge. The existence of such an anti-particle, the positron, was later confirmed by observation [10].

The correctly formulated Lagrangian requires a method for multiplying two spinors in order to obtain a Lorentz scalar. This is done by introducing the adjoint field  $\bar{\psi}(x) \equiv \psi^\dagger \gamma^0$  and the Lorentz invariant Lagrangian can be written as

$$\mathcal{L}_{\text{Dir.}} = \bar{\psi}(x)(i\gamma^\mu \partial_\mu - m)\psi(x) \quad (2.6)$$

Like in the case of scalar fields, it is also convenient to Fourier expand the Dirac field. A fundamental difference is that the fields as well as the creation and annihilation operators now obey the *anticommutator* algebra, while the corresponding particles obey the Fermi-Dirac statistics.

The electric and magnetic fields are classically described by the Maxwell equations. These can be simplified by introducing the four-component vector field  $A_\mu = (\phi, \vec{A})$ , where  $\phi$  is the scalar potential and  $\vec{A}$  the vector potential. The Maxwell equations can then be simplified to two equations where the  $\vec{E}$  and  $\vec{B}$  fields are expressed in terms of  $A_\mu$ . These fields appear then to be elements of an antisymmetric field strength tensor  $F_{\mu\nu}$

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix}$$

$$F_{\mu\nu} = -F_{\nu\mu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (2.7)$$

Generalizing the case to include also massive vector fields one gets the Lagrangian density

$$\mathcal{L}_{\text{Vec.}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}m^2 A^\mu A_\mu \quad (2.8)$$

which yields with the help of the Euler-Lagrange equation (Eq. 2.4) the equation of motion for massive vector particles, the so called Proca equation:

$$(\partial_\nu \partial^\nu - m^2)A_\mu = 0 \quad (2.9)$$

## 2.2 The gauge principle

Some of the mathematical corner stones of the field theory of particle physics are based on Noether's theorem which relates symmetries of the action under a group of transformations with conserved quantities, or in other words, constants of motion. However, the principle of symmetries is even more profound as it actually also implies the dynamics of the theory. The theory of *local gauge transformations* as a foundation of the SM was developed by Salam and Ward [11] who wrote in 1959:

*Our basic postulate is that it should be possible to generate strong, weak and electromagnetic interaction terms (with all their correct symmetry properties and also with clues regarding their relative strengths) by making local gauge transformations on the kinetic energy terms in the free Lagrangian for all particles.*

Quantum Electro Dynamics (QED) is an example of such a gauge theory. Starting from the Dirac Lagrangian (Eq. 2.6) for charged half-integer spin particles we require it to be invariant under infinitesimal local ( $\vec{x}$ -dependent)  $U(1)$  transformations:

$$e^{i\alpha(x)} = 1 + i\alpha(x) + O(\alpha^2) \quad (2.10)$$

such that the field and its derivative transform as

$$\psi(x) \rightarrow e^{i\alpha(x)}\psi(x), \quad \partial_\mu \psi(x) \rightarrow e^{i\alpha(x)}\partial_\mu \psi(x) + i\partial_\mu \alpha(x)e^{i\alpha(x)}\psi(x). \quad (2.11)$$

However, the last term with  $\partial_\mu \alpha$  spoils the invariance. The invariance is re-established by introducing the covariant derivative which transforms in the same way as the field

$$D_\mu \psi(x) \rightarrow e^{i\alpha(x)}D_\mu \psi(x). \quad (2.12)$$

The covariant derivative is constructed by adding a vector field, represented by the  $A_\mu$ -term, which is a so-called gauge field, following the minimal substitution

$$\partial_\mu \rightarrow D_\mu \equiv \partial_\mu + ieA_\mu. \quad (2.13)$$

Now we require the derivative to transform as the field in order to preserve the invariance of the Lagrangian:

$$\begin{aligned} D_\mu \psi &= (\partial_\mu + ieA_\mu)\psi \\ &\rightarrow (\partial_\mu + ieA'_\mu)e^{i\alpha}\psi \\ &= e^{i\alpha}(\partial_\mu + i\partial_\mu \alpha + ieA'_\mu)\psi \end{aligned}$$

In order to obtain that we require the gauge field to transform in an appropriate way

$$A' = A_\mu - \frac{1}{e}\partial_\mu\alpha. \quad (2.14)$$

Replacing  $A'$  with the equation above, we obtain the correct transformation of the covariant derivative. The resulting gauge invariant Lagrangian density with the additional free gauge field part expressed in terms of the field strength tensor  $F_{\mu\nu}$  is now the familiar QED Lagrangian:

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi - e\bar{\psi}\gamma^\mu A_\mu\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \quad (2.15)$$

The “missing” part in this equation compared to the Lagrangian of a massive vector field given by Eq. 2.8, is the mass term proportional to  $A^\mu A_\mu$ . It cannot be added by hand as it would ruin the gauge invariance since it transforms like

$$A^\mu A_\mu \rightarrow (A^\mu - \frac{1}{e}\partial^\mu\alpha)(A_\mu - \frac{1}{e}\partial_\mu\alpha) \neq A^\mu A_\mu. \quad (2.16)$$

The missing mass term is not problematic in the case of QED as the photon is massless. However, it becomes a problem when dealing with the massive intermediate vector bosons in weak interactions and the problem requires a different solution which will be described later.

In order to prepare the grounds for later presentation of the theory it is worth mentioning that the field strength tensor  $F_{\mu\nu}$  can be expressed in terms of the commutator of the two covariant derivatives

$$F_{\mu\nu} = -\frac{1}{e}[D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (2.17)$$

This shows that the interactions are related to the commutator relations of the transformation group generators. In QED which is a simple Abelian U(1) case, we see that the photons do not have self interaction terms.

## 2.3 Non-abelian gauge theory

The gauge procedure from the previous section can also be applied to non-Abelian group transformations. These are groups where the generators  $T_a$  do not commute but obey the Lie algebra

$$[T_a, T_b] = i\epsilon_{abc}T_c \quad (2.18)$$

where the  $\epsilon_{abc}$  are the structure constants of the group.

One common example is related to the isospin transformations. An isodoublet can describe a u-type quark and a d-type quark with the third isospin component  $+1/2$  and  $-1/2$ , respectively. One could also use as example a doublet of left-handed leptons, one neutral ( $\nu$ ) and one charged ( $e^-$ ), or a proton and a neutron. Their wave functions are invariant under a transformation  $U$

$$\psi \rightarrow \psi' = U\psi \quad (2.19)$$

where  $U$  is unitary ( $U^\dagger U = UU^\dagger$ ) and where  $\det(U) = 1$ . One such group is labelled  $SU(2)$  and its generators are known as the Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Multiplied by a factor  $1/2$  they represent the generators of rotations of spin  $1/2$ -particles,  $\vec{T} = \vec{\sigma}/2$ . A general rotation requires also gauge parameters  $\alpha^a$  with  $a = 1, 2, 3$ . We let the gauge transformation be local by making it  $\vec{x}$ -dependent and obtain the infinitesimal transformation

$$\psi(x) \rightarrow e^{i\alpha^a(\vec{x})T^a} \psi(x) \simeq (1 + i\alpha^a(\vec{x})T^a)\psi(x) \quad (2.20)$$

In order to simplify the notation we omit the index  $a$  and write the transformation as vector in  $a$ :  $\vec{U}(\alpha) \equiv e^{i\alpha^a(x)T^a}$ . The derivative transforms now as

$$\partial_\mu \psi(x) \rightarrow \vec{U}(\alpha) \partial_\mu \psi(x) + (\partial_\mu \vec{U}(\alpha)) \psi(x) \quad (2.21)$$

Again, we introduce the covariant derivative

$$D_\mu \equiv (\partial_\mu + igT^a A_\mu^a) \quad (2.22)$$

via the minimal coupling to vector fields with a generalized charge  $g$  equivalent to the electric charge in Eq. 2.13. The number of introduced vector fields equals the number of generators of the group. Since the Lagrangian is a function of the field and its derivative, we require also the covariant derivative to transform as the field. For convenience we write the derivative as a vector in  $a$  with the gauge field  $\vec{W}_\mu = T^a A_\mu^a$  and obtain

$$\vec{D}_\mu \psi \rightarrow \vec{U}(\alpha) \vec{D}_\mu \psi = \vec{U}(\alpha) \partial_\mu \psi + (\partial_\mu \vec{U}(\alpha)) \psi - ig \vec{W}'_\mu \vec{U}(\alpha) \psi. \quad (2.23)$$

We require then that the gauge field  $\vec{W}_\mu$  transforms in a way that gives us the correct transformation of the derivative. This implies that

$$\vec{W}'_\mu = \vec{U}(\alpha) \vec{W}_\mu U^{-1}(\alpha) - \frac{i}{g} (\partial_\mu \vec{U}(\alpha)) \vec{U}^{-1}(\alpha) \quad (2.24)$$

or in the infinitesimal form where we have used the commutator relation

$$A'^a_\mu = A^a_\mu - \epsilon_{abc} \alpha^b(x) A^c_\mu - \frac{1}{g} \partial_\mu \alpha^a(x) \quad (2.25)$$

In addition we add the kinetic term for the free field which we express using the commutator of the covariant derivatives

$$\vec{F}_{\mu\nu} = -\frac{i}{g} [\vec{D}_\mu, \vec{D}_\nu]. \quad (2.26)$$

Due to the non-abelian nature of the transformations it becomes

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g\epsilon_{abc} A^b_\mu A^c_\nu. \quad (2.27)$$

Finally we can write down the gauge invariant Lagrangian

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi - g(\bar{\psi}\gamma^\mu T_a \psi) A^a_\mu - \frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu}. \quad (2.28)$$

The crucial difference compared to the Abelian case is that in the non-Abelian theory, the free field Lagrangian contains also self interactions involving the different gauge fields represented by  $A$ :

$$\mathcal{L}_A \propto (\partial A - \partial A)^2 + g(\partial A - \partial A)AA + g^2AAAA \quad (2.29)$$

$$\text{propagator} \quad \text{triple-} \quad \text{quartic-} \quad \text{coupling} \quad (2.30)$$

Like in the Abelian case, the Lagrangian does not contain any mass term for the gauge field. We know that it should be proportional to  $A_\mu^a A^{a\mu}$ , but adding it by hand will destroy the gauge invariance as we would obtain terms similar to the expression in Eq. 2.16.

Deriving the Lagrangian density for the non-Abelian theory, SU(2) group was used as an example. The three gauge fields represent the vector bosons, two charged and one neutral. However, the actual particle content will only become clear after the spontaneous symmetry breaking of the unified electroweak SU(2)×U(1) group which will be described in Section 2.4.

The other non-Abelian group of the SM is the SU(3) group of the phase transformations on the quark color fields and which thus lies the foundation of the gauge theory of Quantum Chromo Dynamics (QCD). Quarks are fermions described by the Dirac spinors  $q(x)$ . The starting point of QCD is the Lagrangian density found in Eq. 2.6, however the fields are now triplets related to the three color states. The local unitary gauge transformation is as before defined as

$$U = e^{\alpha^a(x)T^a} \quad (2.31)$$

The generators  $T^a$  obey the commutator algebra of Eq. 2.18 but in case of QCD they are defined as:  $T^a = \lambda^a/2$  and the structure constants are the more complex  $f_{abc}$ . The  $\lambda$  matrices refer to the eight  $3 \times 3$  Gell-Mann matrices [12]. In order to preserve the gauge invariance of the Lagrangian we introduce as many vector fields as there are generators in the group, so in QCD there are eight fields and  $a = 1, 2, \dots, 8$ . The resulting Lagrangian is as follows:

$$\mathcal{L} = \bar{q}(i\gamma^\mu \partial_\mu - m)q - g_s(\bar{q}\gamma^\mu T_a q)G_\mu^a - \frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu} \quad (2.32)$$

The  $G_{\mu\nu}$  is the gluonic field strength tensor and the complete expression is a sum over all quark flavors. Compared with Eq. 2.28 it differs only by the coupling constant  $g_s$ , quark triplets  $q$  and the generators  $T_a$  which are now  $3 \times 3$  matrices. The gauge particle of the strong force is the so called gluon. It is interesting to note that gluons like photons are massless, but they do carry a color charge. As shown in Eq. 2.29, due to the non-Abelian nature of the group, they have both triple and quartic couplings where the strength of the interactions is regulated by the QCD coupling constant  $g_s$ .

Although the gauge transformations lead to correct Lagrangian densities, the gauge fields require some additional constraints in order to be able to formulate the propagators. This procedure is called gauge fixing and the resulting propagators are ready to be applied in the calculations of physical quantities like cross sections following the so-called Feynman rules.

## 2.4 The unbearable lightness of gauge bosons

Throughout the derivations of the gauge theories in the previous sections we have seen that the gauge bosons of SU(2) remain massless. Studies of the weak interactions show



however, that this is not the case, but the mass has to be introduced in a way that preserves the gauge invariance. The solution has become known as the Higgs mechanism [13–17] after one of its developers. The mechanism of spontaneous symmetry breaking not only proposes a solution to the mass problem, it is also an essential condition for making the theory of weak interactions renormalizable as it was shown years later by t’Hooft [18].

The Higgs mechanism and finally the full Glashow-Weinberg-Salam model [19] known as the full Standard Model will be introduced in the following steps showing:

- Spontaneous symmetry breaking (SSB) and the Goldstone boson [20]
- SSB in the context of local Abelian gauge theory
- SSB in the context of local non-Abelian theory – the Higgs mechanism
- Formulation of the  $SU(2)_L \times U(1)_Y$  theory of electroweak interaction and applying the earlier results to formulate the full SM Lagrangian

The introduction will show some key steps of the derivation. Although several sources have been used in preparation of the section, for the sake of consistency, the presentation will follow the notation in [4].

## Spontaneously broken global gauge symmetry – Goldstone boson

The Higgs mechanism requires that we introduce a scalar field and in order to make it suitable for our purpose, it has to be a complex field of the form:

$$\phi = \frac{\phi_1 + i\phi_2}{\sqrt{2}} \quad (2.33)$$

It is described by the following Lagrangian

$$\mathcal{L} = (\partial_\mu \phi^*)(\partial^\mu \phi) - V(\phi) = (\partial_\mu \phi^*)(\partial^\mu \phi) - \mu^2 \phi^* \phi - \lambda(\phi^* \phi)^2 \quad (2.34)$$

(with  $\lambda > 0$ ) which is invariant under the global  $U(1)$  symmetry. The minimum of the potential  $V(\phi)$  represents the ground state of the vacuum. If  $\mu^2 > 0$  the Lagrangian simply describes a self interacting scalar field with mass  $\mu$ . However,  $\mu^2$  can be negative and this results in the famous “Mexican hat” potential shown in Figure 2.1 with a minimum along a circle with radius squared given by:

$$\phi_1^2 + \phi_2^2 = v^2 \quad \text{where} \quad v^2 = -\frac{\mu^2}{\lambda}. \quad (2.35)$$

Changing the sign of  $\mu^2$  we no longer have a minimum in  $\phi = 0$  but infinitely many minima along the circle with radius  $v$ . The field is then translated to the minimum valley and at the same time we choose a particular point by setting  $\phi_1 = v$  and  $\phi_2 = 0$ . The field is then expanded around the chosen minimum in terms of the fields  $\eta(x)$ , in the radial direction, and  $\xi(x)$  in the tangential direction:

$$\phi(x) = \frac{1}{\sqrt{2}} [v + \eta(x) + i\xi(x)] . \quad (2.36)$$

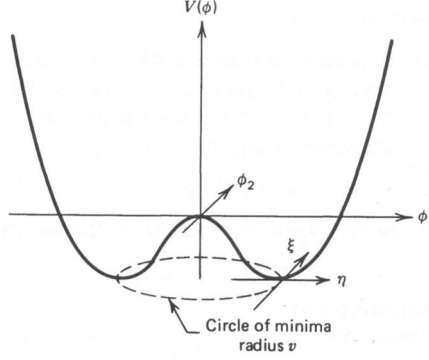


Figure 2.1: The Higgs potential  $V(\phi)$  for a complex scalar field where  $\mu^2 < 0$  and  $\lambda > 0$ . [4]

Inserting the new  $\phi(x)$  into the Lagrangian given by 2.34, we obtain

$$\mathcal{L}' = \frac{1}{2}(\partial_\mu \xi)^2 + \frac{1}{2}(\partial_\mu \eta)^2 + \mu^2 \eta^2 + \text{const.} + \text{interaction terms in } \eta, \xi. \quad (2.37)$$

We now see that a mass term of the form  $-\frac{1}{2}m_\eta^2 \eta^2$  has appeared with  $m_\eta = \sqrt{-2\mu^2}$ . By choosing a particular ground state and expanding the  $\mathcal{L}$  about this stable minimum we have introduced a mass term which does not violate the gauge invariance. However, the  $\xi$  field remains massless and is known as the Goldstone boson [20]. Intuitively we get an idea behind the phenomenon noting that the potential in the tangent  $\xi$ -direction is flat, while the  $\eta$ -field is an excitation up the slope of the potential. The Goldstone theorem generalizes the result stating that a massless scalar occurs whenever a continuous symmetry of a physical system is “spontaneously” broken. Although this is not the whole answer to the problem of mass, it is the foundation for the next step which combines the spontaneous symmetry breaking with the requirement of Lagrangian invariance under local gauge transformation.

## Spontaneous symmetry breaking of local gauge symmetry – Higgs mechanism

The decisive step is now to promote the gauge transformation to a local one

$$\phi \rightarrow \phi' = e^{i\alpha(x)} \phi \quad (2.38)$$

In Section 2.2 we saw how to obtain the gauge invariant Lagrangian by introducing the necessary gauge fields and a covariant derivative  $D_\mu$  which included these fields and replaced the normal derivative. Requiring that the new derivative term  $D_\mu \phi$  transforms as the field and choosing a field transformation that ensures this, we obtain, starting from Eq. 2.34, the following Lagrangian:

$$\mathcal{L} = (\partial^\mu + ieA_\mu)\phi^*(\partial_\mu - ieA_\mu)\phi - \mu^2 \phi^* \phi - \lambda(\phi^* \phi)^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (2.39)$$

where the last term accounts for the free gauge field. In order to generate spontaneous symmetry breaking we concentrate on the case where  $\lambda > 0$  and  $\mu^2 < 0$ . Like in the

previous section, we break the symmetry by choosing a particular vacuum and insert into the Lagrangian the field given by Eq. 2.36 and obtain

$$\mathcal{L}' = \frac{1}{2}(\partial_\mu \xi)^2 + \frac{1}{2}(\partial_\mu \eta)^2 - v^2 \lambda \eta^2 + \frac{1}{2}e^2 v^2 A_\mu A^\mu - ev A_\mu \partial^\mu \xi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \text{int. terms.} \quad (2.40)$$

The new Lagrangian contains a massless Goldstone boson  $\xi$ , a massive scalar particle  $\eta$  and finally a mass term for the gauge field proportional to  $A_\mu A^\mu$  ( $A_\mu$  represents here an arbitrary field). In summary the masses are

$$m_\xi = 0 \quad m_\eta = \sqrt{2\lambda v^2} \quad m_A = ev. \quad (2.41)$$

The only remaining problem is the bilinear  $A_\mu \partial^\mu \xi$ -term which indicates that the longitudinal component of the gauge field couples to the Goldstone boson with the strength equal to  $m_A$ . We express the scalar field in terms of a new set of real fields:

$$\phi \rightarrow \sqrt{\frac{1}{2}}(v + h(x))e^{i\phi(x)/v} \quad (2.42)$$

and implement the unitary gauge where the physical field is the following superposition

$$A_\mu \rightarrow A_\mu + \frac{1}{ev}\partial_\mu \theta(x). \quad (2.43)$$

With this substitution the vector field acquires a longitudinal degree of freedom from the Goldstone boson. The final Lagrangian is independent of the  $\theta$  field and expressed in terms of the real  $h$ -field it becomes:

$$\mathcal{L}'' = \frac{1}{2}(\partial_\mu h)^2 - v^2 \lambda h^2 + \frac{1}{2}e^2 v^2 A_\mu A^\mu - \lambda v h^3 - \frac{1}{4}\lambda h^4 + \frac{1}{2}e^2 A_\mu^2 h^2 + v e^2 A_\mu A^\mu h - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \quad (2.44)$$

$\mathcal{L}''$  describes now only two fields: a massive vector boson  $A_\mu$  and a massive scalar  $h$ . The procedure by which we arrived at this result is known as the Higgs mechanism.

## The broken local SU(2) symmetry

On our way towards the full SM Lagrangian we now apply the procedure from the previous section on the local SU(2) gauge theory. The fields are now represented by doublets and the complex conjugation in the scalar Lagrangian given by Eq. 2.34 is substituted by Hermitian conjugation. The  $\phi$  field is taken to be a doublet of complex scalar fields

$$\phi = \begin{pmatrix} \phi_\alpha \\ \phi_\beta \end{pmatrix} = \sqrt{\frac{1}{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \quad (2.45)$$

In order to establish a Lagrangian which is invariant under the local SU(2) transformation we follow the derivation outlined in Section 2.3 introducing infinitesimal transformation of the field and the covariant derivative

$$\phi(x) \rightarrow \phi'(x) = (1 + i\vec{\alpha}(x) \cdot \vec{\tau}/2)\phi(x), \quad D_\mu = \partial_\mu + ig\frac{\vec{\tau}_a}{2}W_\mu^a. \quad (2.46)$$

We require the field to transform in the following way

$$\vec{W}_\mu \rightarrow \vec{W}'_\mu = \vec{W}_\mu - \frac{1}{g}\partial_\mu \vec{\alpha} - \vec{\alpha} \times \vec{W}. \quad (2.47)$$

and obtain the familiar gauge invariant Lagrangian

$$\mathcal{L} = \left( \partial_\mu \phi + ig \frac{\vec{\tau}}{2} \cdot \vec{W}_\mu \phi \right)^\dagger \left( \partial^\mu \phi + ig \frac{\vec{\tau}}{2} \cdot \vec{W}^\mu \phi \right) - V(\phi) - \frac{1}{4} \vec{W}_{\mu\nu} \vec{W}^{\mu\nu} \quad (2.48)$$

The field strength tensor in the last term is of the same form as we found for the non-Abelian case given by Eq. 2.27. So far everything is familiar. We now look at the interesting case where  $\lambda > 0$  and  $\mu^2 < 0$  and the potential obtains a minimum when

$$\phi^\dagger \phi = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = -\frac{\mu^2}{2\lambda}. \quad (2.49)$$

In order to spontaneously break the symmetry we choose a particular ground state

$$\phi_1 = \phi_2 = \phi_4 = 0, \quad \phi_3^2 = -\frac{\mu^2}{\lambda} = v^2. \quad (2.50)$$

Following a similar argument as when setting the ground state for the  $U(1)$  case and applying unitary gauge, the scalar field is now described by the vacuum expectation value and the expansion about:

$$\phi(x) = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \quad (2.51)$$

The mass terms are found when we insert this field in the Lagrangian Eq. 2.48. In order to be able to see the structure of the mass terms it is enough to insert the ground state, that means the field in Eq. 2.51 without the  $h(x)$  field. Concentrating on the first term of Eq. 2.48 we find

$$\left| ig \frac{\vec{\tau}}{2} \cdot \vec{W}_\mu \phi \right|^2 = \frac{g^2 v^2}{8} [W_\mu^1 W^{1\mu} + W_\mu^2 W^{2\mu} + W_\mu^3 W^{3\mu}] \quad (2.52)$$

which is clearly the boson mass of the form  $\frac{1}{2} M_A A_\mu A^\mu$ . The new Lagrangian will thus contain three massive gauge boson fields, which have acquired mass thanks to the Goldstone bosons, as well as a massive Higgs particle.

### 2.4.1 Building the full Standard Model

We now proceed towards the formulation of the full SM Lagrangian. In order to accommodate both the weak and electromagnetic interactions in one theory Glashow [21] proposed in 1961 the  $SU(2) \times U(1)$  gauge group as the foundation of such a theory. The weak interactions couple to the left-handed fermions so it suggested that the theory should be formulated in terms of such helicity eigenstates and that the  $SU(2)$  group should be associated with the weak isospin. The  $U(1)$  group was associated with the hypercharge  $Y$ . These two, the weak isospin and hypercharge, relate to each other through the Gell-Mann – Nishijima formula for the electric charge:  $Q = T_3 + Y/2$  where  $T_3$  is the third component of the isospin. The combined group requires three gauge fields for the  $SU(2)$  group ( $W^1, W^2, W^3$ ) and a neutral field ( $B$ ) related to the  $U(1)$  group.

In the following text the components of the theory will not be derived (see [4] or [8] for further details). However, building on the derivations from the previous sections, it will rather be a summary of the bits and pieces required to generate the masses and finally formulate the Glashow-Weinberg-Salam theory [19].

The formalism of the weak interactions requires that we formulate the Lagrangian in terms of left-handed and right-handed fermion fields. For leptons these are defined as

$$\chi_L = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \quad T = \frac{1}{2}, Y = -1 \quad (2.53)$$

$$\psi_R = e_R^-, \quad T = 0, Y = -2 \quad (2.54)$$

and for quarks as

$$\chi_L = \begin{pmatrix} u \\ d \end{pmatrix}_L, \quad \psi_R = u_R \text{ or } d_R. \quad (2.55)$$

In case of quarks there are also right-handed components for the up-type quarks as they, unlike the neutrinos, have a mass<sup>1</sup>. The original Lagrangian contains two basic interactions: the weak current that couples to the vector field  $\vec{W}_\mu$

$$-ig\vec{J}_\mu \cdot \vec{W}_\mu = -ig\bar{\chi}_L \gamma_\mu \vec{T} \cdot \vec{W}^\mu \chi_L \quad (2.56)$$

and the weak hyper charge current that couples to the  $B$  field

$$-ig'j_\mu^Y B^\mu = -ig'\bar{\psi}\gamma_\mu \frac{Y}{2}\psi B^\mu \quad (2.57)$$

$\vec{T}$  and  $Y$  are the generators of the  $SU(2)_L$  and  $U(1)_Y$  groups respectively and the corresponding gauge transformations of the fields are

$$\chi_L \rightarrow \chi'_L = e^{i\vec{\alpha}(x) \cdot \vec{T} + i\beta(x)Y} \chi_L \quad (2.58)$$

$$\psi_R \rightarrow \psi'_R = e^{i\beta(x)Y} \psi_R \quad (2.59)$$

The generators of the gauge group satisfy the Gell-Mann–Nishijima formula. The electromagnetic interaction is integrated in the expressions 2.56 and 2.57 and can be written as a combination of gauge fields

$$j_\mu^{em} = J_\mu^3 + \frac{1}{2}j_\mu^Y \quad (2.60)$$

such that the physical fields  $A_\mu$  and  $Z_\mu$  representing  $\gamma$  and  $Z$  are orthogonal combinations of the gauge fields  $B_\mu$  and  $W_\mu^3$

$$\begin{aligned} A_\mu &= B_\mu \cos \theta_W + W_\mu^3 \sin \theta_W \\ Z_\mu &= -B_\mu \sin \theta_W + W_\mu^3 \cos \theta_W \end{aligned} \quad (2.61)$$

where the angle  $\theta_W$  is the so called Weinberg mixing angle. Expressed in terms of the physical fields, the neutral current can be written as

$$-igJ_\mu^3 W^{3\mu} - ig'j_\mu^Y B^\mu = -iej_\mu^{em} A^\mu - \frac{ie}{\sin \theta_W \cos \theta_W} [J_\mu^3 - \sin^2 \theta_W j_\mu^{em}] Z^\mu \quad (2.62)$$

From this formula we can find a relation between the coupling constants  $g$  and  $g'$ , the electrical charge  $e$  and the angle  $\theta_W$ :

$$e = g \sin \theta_W = g' \cos \theta_W \quad (2.63)$$

---

<sup>1</sup>There are currently strong indications that neutrinos do have mass, however the mechanism behind it is beyond the scope of this brief summary [22].

Before arriving to the breaking of the  $SU(2)_L \times U(1)_Y$  gauge symmetry we write down the corresponding Lagrangian using the electron and neutrino as example and insert the correct values for the hypercharge:

$$\begin{aligned} \mathcal{L}_1 = & \bar{\chi}_L \gamma \left[ i\partial_\mu - g\frac{\vec{\tau}}{2} \cdot \vec{W}_\mu - g'(-\frac{1}{2}B_\mu) \right] \chi_L \\ & + \bar{e}_R \gamma^\mu [i\partial_\mu - g'(-1)B_\mu] e_R - \frac{1}{4} \vec{W}_{\mu\nu} \cdot \vec{W}^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \end{aligned} \quad (2.64)$$

This lengthy formula involves many of the earlier derived components and its structure with kinetic terms and free fields, is the same for the three families of leptons and quarks grouped in helicity doublets and singlets. However, in this representation all the fields are massless, not only the gauge fields, but also the fermions for which the usual mass-term  $-m\bar{\psi}\psi$  mixes the left- and right-handed fields and ruins the gauge invariance.

In order to apply the Higgs mechanism one has to add a  $SU(2) \times U(1)$  scalar field Lagrangian

$$\mathcal{L}_2 = \left| \left( i\partial_\mu - g\vec{T} \cdot \vec{W}_\mu - g'(-\frac{1}{2}B_\mu) \right) \phi \right|^2 - V(\phi) \quad (2.65)$$

The simplest scalar field that can be introduced with hypercharge 1 is

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad (2.66)$$

where

$$\phi^+ = (\phi_1 + i\phi_2)/\sqrt{2} \quad \phi^0 = (\phi_3 + i\phi_4)/\sqrt{2} \quad (2.67)$$

The potential  $V(\phi)$  is the same as in Eq. 2.34 and again the relevant case is when  $\lambda > 0$  and  $\mu^2 < 0$  and we choose the same ground state as Eq. 2.50. Inserting it into the Lagrangian given by Eq. 2.65 leads to the following gauge boson mass terms:

$$\begin{aligned} \left| \left( -g\vec{T} \cdot \vec{W}_\mu - g'(-\frac{1}{2}B_\mu) \right) \phi \right|^2 = & \left( \frac{1}{2}vg \right)^2 W_\mu^+ W^{-\mu} \\ & + \frac{1}{8}v^2 (W_\mu^3, B_\mu) \begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix} \begin{pmatrix} W^{3\mu} \\ B^\mu \end{pmatrix}. \end{aligned} \quad (2.68)$$

We have written the charged fields as

$$W^\pm = (W^1 \mp iW^2)/\sqrt{2}. \quad (2.69)$$

The first term on the right-hand side of Eq. 2.68 provides the  $W$ -mass

$$M_W = \frac{1}{2}vg. \quad (2.70)$$

The second term on the right-hand side gives the mass for the photon and the  $Z$ . By diagonalizing the matrix and using the physical fields from Eq. 2.61 gives the following

$$A_\mu = \frac{g'W_\mu^3 + gB_\mu}{\sqrt{g^2 + g'^2}} \quad \text{and} \quad M_A = 0 \quad (2.71)$$

$$Z_\mu = \frac{gW_\mu^3 - g'B_\mu}{\sqrt{g^2 + g'^2}} \quad \text{and} \quad M_Z = \frac{1}{2}v\sqrt{g^2 + g'^2} \quad (2.72)$$

The relation between the mass of the  $Z$  and the  $W$  (using Eq. 2.63 and Eq. 2.70) is

$$\frac{M_W}{M_Z} = \cos \theta_W. \quad (2.73)$$

Now that the bosons have acquired mass through the Higgs mechanism, it still remains to formulate the fermion mass generation. It is done with the help of the same Higgs <sup>2</sup> doublet that was chosen for the bosons in Eq. 2.66 and the appropriate Lagrangian is

$$\mathcal{L}_3 = -G_e \left[ (\bar{\nu}_e, \bar{e})_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R + \bar{e}_R (\phi^-, \bar{\phi}^0) \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \right] \quad (2.74)$$

$G_e$  is the electron Yukawa coupling. Spontaneously breaking the symmetry by inserting the ground state Eq. 2.51 we arrive at the required Lagrangian

$$\mathcal{L}_3 = -m_e \bar{e}e - \frac{m_e}{v} \bar{e}e h \quad \text{where} \quad m_e = \frac{G_e v}{\sqrt{2}}, \quad m_\nu = 0. \quad (2.75)$$

In the second term of  $\mathcal{L}_3$  we see that the fermions, here represented by the electron, have a coupling to the Higgs,  $G_e$ , which is proportional to the mass.

Giving mass to quarks follows basically the same approach, although there are two complications. The Higgs field used in Eq. 2.74 is only capable of giving mass to the down-type quarks. In order to give mass to the up-type quarks, one has to introduce

$$\phi_c = -i\tau_2 \phi^* = \begin{pmatrix} -\bar{\phi}^0 \\ \phi^- \end{pmatrix} \quad (2.76)$$

which upon symmetry breaking acquires non-zero entries in the upper row. The second complication is that the weak interaction operates on the doublets  $(u, d')_L, (c, s')_L, \dots$  where the primed fields are orthogonal combinations of the physical mass eigenstates with mixing given by the Cabibbo angle  $\theta_C$  and two additional angles (see Section 2.5). The resulting Lagrangian follows the pattern of the one for leptons

$$\mathcal{L}_4 = -G_d^{ij} (\bar{u}_i, \bar{d}'_i)_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} d_{jR} - G_u^{ij} (\bar{u}_i, \bar{d}'_i)_L \begin{pmatrix} -\bar{\phi}^0 \\ \phi^- \end{pmatrix} u_{jR} + \text{h.c.} \quad (2.77)$$

The Yukawa coupling  $G_{u/d}^{ij}$  is now a matrix.

After breaking the symmetry one obtains similar mass terms as in Eq. 2.75, now proportional to  $\bar{d}_i d_i$  where  $i$  represents the family and the interaction with the Higgs field is proportional to the mass.

Although the mechanism for particle mass generation is established, it does not predict any of the coupling constants  $G$ . Neither is the fundamental choice of a negative  $\mu^2$  factor fully explained. The Higgs particle must still be experimentally confirmed and its mass measured. It was not found at LEP, but a lower limit of 114.4 GeV was set at 95% confidence level (C.L.) [23]. Combining this result with precision electroweak measurements sets an upper limit at 186 GeV at the 95% C.L. [24].

Recently the Tevatron presented combined SM Higgs limits excluding the mass range between 162 GeV and 166 GeV at the 95% C.L. [25]. Figure 2.2 (left) shows the Tevatron exclusion plot where  $R_{lim}$  is the ratio of the limits obtained to the rate predicted by SM. The right panel shows the combined LEP  $\chi^2$  vs.  $m_H$  showing in yellow/gray both the lower limit and the recent Higgs mass range excluded by the Tevatron.

---

<sup>2</sup> $\phi^-$  and  $\bar{\phi}^0$  are the complex conjugated fields from Eq. 2.67.

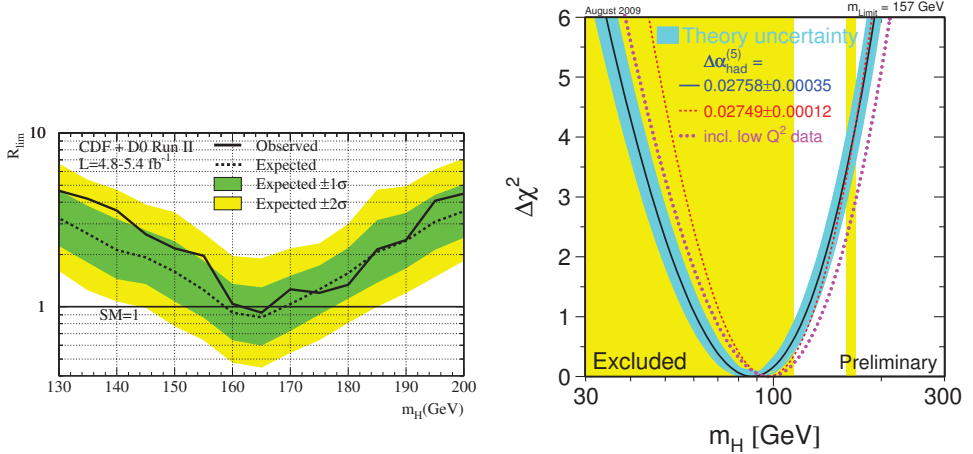


Figure 2.2: Left: Tevatron combined results.  $R_{lim}$  is the ratio of the limits obtained to the rate predicted by SM [25]. Right:  $\Delta\chi^2$  vs.  $m_H$  indicating also the Tevatron excluded region [24] (160-170 GeV [26]).

## 2.5 Weak interactions and the CKM matrix

Let us for a moment return to the primed quark fields. Observing processes like  $K^+ \rightarrow \mu^+ \nu_\mu$  it became clear that there must be a weak charged current that couples  $u$  and  $\bar{s}$  quarks which belong to different families. It was therefore assumed that it couples to “rotated” quark states which can be written as orthogonal combinations of the mass eigenstates. Generalizing the formulation to include all three families it can be expressed by a unitary matrix  $U$  defined by the flavor-changing weak processes and the weak eigenstate as ( $i$  and  $j$  are flavor indices)

$$d_L^i = \sum_j U_d^{ij} u_L^j. \quad (2.78)$$

The neutral current which is proportional to terms like  $\bar{u}_i' \gamma^\mu u_i'$  are not affected by this change of basis. While the charged current is

$$W^+ \bar{u}_L^i \gamma^\mu d_L^i + W^- \bar{d}_L^i \gamma^\mu u_L^i = W^+ \bar{u}_L^i \gamma^\mu (U_u^{\dagger ik} U_d^{kj}) d_L^j + W^- \bar{d}_L^i \gamma^\mu (U_u^{\dagger ik} U_d^{kj}) u_L^j. \quad (2.79)$$

The matrix in the parenthesis on the right hand side is the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix [27] [28] where the elements are

$$V_{CKM}^{ij} = U_u^{\dagger ik} U_d^{kj} \quad (2.80)$$

Written out it becomes

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}$$

The elements along the diagonal are largest, close to unity, as they represent transitions within one family, while the off-diagonal terms represent mixing between the families and are considerably smaller.



Given that the CKM matrix is a unitary  $3 \times 3$  matrix, it can be parametrized in terms of three real parameters and one complex phase factor which introduces CP-violation. The phenomenon of CP-violation was first discovered in studies of decays of neutral kaons and is currently extensively studied in decays of  $B$  mesons.

One of the reasons why CP-violation is so important is that it states that matter and anti-matter is somehow different. According to basic assumptions, there should have been produced equal amounts of matter and anti-matter in the Big Bang which later would simply “cancel out”. While this is not the case, CP-violation is one of the keys to the understanding of this matter-antimatter asymmetry. However, the CP-violation introduced by the CKM matrix is not large enough to give the observed matter-antimatter asymmetry.



# Chapter 3

## Beyond the Standard Model

The SM has so far been a remarkably successful theory. It describes very well the current high precision measurements and gives bold predictions like cross sections and branching fractions. The top mass could for example be predicted several years before it was discovered. Nevertheless, there are many reasons for believing the SM is not the final theory. An ideological reason is that the SM contains a large number of parameters which have to be set by experiments, while one would prefer the theory itself to predict most of them.

John Ellis associates issues beyond the Standard Model into three groups: **Flavor**, **Unification** and **Mass** [29]. As we will see, the Supersymmetric - SUSY - extensions of the SM which will be addressed in this section propose solutions to several issues associated with these headings.

The problem of mass is related to the fact that contrary to what the basic QFT Lagrangian show, the particles *do* have a mass which characterises the different types of particles. It is believed that the Higgs mechanism introduced in the previous section is responsible for breaking the electroweak symmetry and giving mass to both the electroweak bosons and to fermions.

The Higgs mechanism predicts a spin-0 Higgs boson which is for many reasons believed to be light. This view is strongly supported by global electroweak fits which suggest that the SM Higgs mass,  $m_H$ , could be around 115 GeV [30, 31].

The mass of the  $W^\pm$  boson is closely related to  $m_H$  and we know it is  $80.398 \pm 0.025$  GeV [32] which is very small compared to the Planck mass  $m_P \sim 10^{19}$  GeV. The Planck mass is considered to be the upper limit of the physics we would like to explain in our theories.

The radiative corrections to the Higgs mass, and thus also to the  $W^\pm$  mass, go like the square of a cut-off parameter  $\Lambda$

$$\delta m_H^2 \sim O\left(\frac{\alpha}{\pi}\right)\Lambda^2. \quad (3.1)$$

If we apply the SM all the way up to the Planck scale the correction would be 36 orders of magnitude greater than the physical mass of these particles [33]. Most physicists think that a good theory should have a mechanism that protects it from such unphysical development.

The solution starts from the observation that corrections from fermion and boson loops like those of Figure 3.1 have opposite signs. Fermion loops have a minus sign, and it is the heavy top quark which is responsible for the largest contribution to  $\delta m_H$ . One could keep the correction small using fine-tuning, but this must be done at a very precise level. Many physicists find it therefore “unnatural” and look for a more straight forward way of

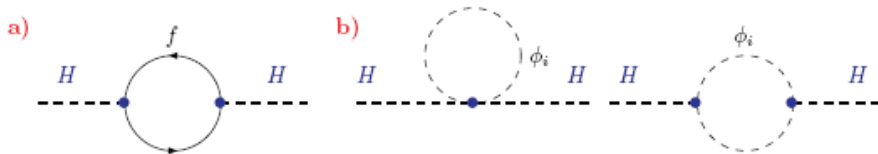


Figure 3.1: Examples of loop diagrams contributing to the Higgs mass. Fermion loop contribution a) is negative while the boson loops b) are positive. [34]

counterbalancing the large corrections. What we observe here is an aspect of the so-called “hierarchy problem”. The SM energy scale is so much lower than the Planck scale, while it is still a criterion for a successful theory that it should be valid all the way to this upper bound. The SM is therefore not considered to be the full solution, but a good low-energy effective theory. The integration limit, the cut-off  $\Lambda$  which drives the mass correction, can be considered as the scale where we would expect new phenomena. The hint from the observation of the opposite signs of the contributions from fermion and boson loops suggests that the new physics should include something bosonic. On the other hand, there are also divergent contributions from SM bosons and from the Higgs itself, and that could be taken care of by something fermionic. An equal number of fermions and bosons in the theory would modify the Higgs mass corrections to become

$$\delta m_H^2 \sim O\left(\frac{\alpha}{\pi}\right) \cdot |m_F^2 - m_B^2| \quad (3.2)$$

which would be smaller than  $m_{H,W}^2$  if  $|m_F^2 - m_B^2| \lesssim 1 \text{ TeV}$ . So if SUSY is the solution, there are good reasons to believe that it is within the reach of the LHC experiments which are designed to cover this crucial energy range up to a few TeV.

This is the motivation for introducing a symmetry between fermions and bosons, and grouping them in chiral super-multiplets. The members of such a multiplet would have the same quantum numbers, in principle the same mass <sup>1</sup>, but different spins. A quick look at the available SM particles and their quantum numbers shows that these alone cannot provide such multiplets. SUSY introduces therefore so-called super-partners to all known particles in the SM, but it does not introduce any new interactions. The usual interactions of the SM with their familiar coupling constants are at work.

$$\begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix} \text{ partners: } \begin{pmatrix} \tilde{\nu}_{eL} \\ \tilde{e}_L \end{pmatrix} \quad \begin{pmatrix} u_L \\ d_L \end{pmatrix} \text{ partners: } \begin{pmatrix} \tilde{u}_L \\ \tilde{d}_L \end{pmatrix}$$

In particle physics we are operating with a generalized notion of charge. The most familiar is the U(1) hyper charge and SU(2) weak charge. They correspond to the generators of the symmetry groups. In SUSY we are also looking for such a generator which operating on a state transforms it into a different member of the super-multiplet.

$$Q|boson\rangle = |fermion\rangle \quad Q|fermion\rangle = |boson\rangle$$

In common notation it is denoted  $Q_a$ , where  $a$  is the spinor component.

$$Q_a|J\rangle = |J \pm \frac{1}{2}\rangle$$

<sup>1</sup>SUSY is a broken symmetry and the masses are different. This will be treated in Section 3.6.1

### 3.1 SUSY chiral fields

The simplest SUSY theory must have both scalar and fermion fields. In QED fermions are described by 4-component Dirac spinors which contain both left- and right-handed fields. Handedness, or chirality is a key issue in SUSY. Several formalisms are available, but this introduction will follow the notation of Aitchison [35] which uses 2-component chiral fields and a complex scalar field. The Dirac spinor

$$\Psi = \begin{pmatrix} \psi \\ \chi \end{pmatrix} \quad (3.3)$$

can be decomposed into  $\psi$ , the right-handed field and  $\chi$ , the left-handed.

As it is convenient to formulate the theory in terms of left-handed fields only, we want to express the right-handed field in terms of left-handed. These fields have the same behaviour under rotation (they both have spin 1/2), but transform in a different way under boost. An infinitesimal Lorentz transformation  $V$  which transforms the right-handed and left-handed fields is defined in the following way:

$$V = (1 + i\vec{\epsilon}\frac{\vec{\sigma}}{2} - \vec{\eta}\frac{\vec{\sigma}}{2}) \quad \psi' = V\psi \quad \chi' = V^{\dagger-1}\chi = V^{-1\dagger}\chi. \quad (3.4)$$

where  $\vec{\epsilon} = (\epsilon_1, \epsilon_2, \epsilon_3)$  are parameters specifying the infinitesimal rotation while  $\vec{\eta} = (\eta_1, \eta_2, \eta_3)$  are three infinitesimal velocities. The  $\vec{\sigma}$  are the Pauli matrices. We realize also that  $\sigma_2\chi^*$  transforms like  $\psi$  and we define:

$$\psi_\chi \equiv i\sigma_2\chi^* \quad (3.5)$$

which allows us to write the Lorentz invariant scalar product as:

$$\psi_\chi^\dagger \chi = \chi^T (-i\sigma_2) \chi. \quad (3.6)$$

The  $-i\sigma_2$  reminds us of charge conjugation which can be associated with the operator  $C_0$

$$C_0 = -1\gamma^2 = \begin{pmatrix} 0 & i\sigma_2 \\ -i\sigma_2 & 0 \end{pmatrix}. \quad (3.7)$$

where  $\gamma^2$  is a Dirac matrix. Putting these things together we can now express a right-handed particle as a charge conjugate of the left handed antiparticle.

$$\Psi^{(p)} = \begin{pmatrix} \psi_R^p \\ \chi_L^p \end{pmatrix} = \begin{pmatrix} i\sigma_2\chi_L^{\bar{p}*} \\ \chi_L^p \end{pmatrix} \quad (3.8)$$

### 3.2 Simple SUSY Lagrangian

The simplest supersymmetric Lagrangian contains a complex scalar field  $\phi$  and a left-handed spinor field  $\chi$

$$\mathcal{L} = \partial_\mu\phi^\dagger\partial_\mu\phi + \chi^\dagger i\bar{\sigma}^\mu\partial_\mu\chi \\ \sigma^\mu = (1, \vec{\sigma}) \quad \bar{\sigma}^\mu = (1, -\vec{\sigma}) \quad (3.9)$$

where the  $\sigma$  matrices are the Pauli-matrices.

Like in other field theories, we would like to keep the Lagrangian invariant under specific transformations, in this case under the SUSY transformations which link a scalar field to a fermion field:

$$\phi \rightarrow \phi' = \phi + \delta_\xi \phi \quad (3.10)$$

where  $\xi$  is a constant left-handed spinor. This means that:

$$\delta_\xi \phi = \text{change in } \phi = \text{parameter} \times \text{left-handed fermion field } \chi \quad (3.11)$$

If  $\phi$  is Lorentz invariant, then so must  $\phi'$ . Following the pattern of the Lorentz invariant product of fermion fields from Eq. 3.6, one can write the change as:

$$\delta_\xi \phi = \xi^T (-i\sigma_2) \chi \quad (3.12)$$

We must also find the quantity  $\delta_\xi \chi$ , the change in the left-handed field, and again following the idea of SUSY we can expect that the form of  $\delta_\xi \chi$  is a product of  $\xi$  and  $\phi$ . A dimension of mass argument is used to derive that  $\chi$  must transform as [35]:

$$\delta_\xi \chi = A [i\sigma^\mu (i\sigma_2 \xi^*) \partial_\mu \phi] \quad (3.13)$$

where  $A$  is a factor that needs to be set. It is required that the transformation leaves the Lagrangian invariant, or at least that the change is a total derivative which can be assumed to vanish at the boundaries of the integration. In order to fulfill that,  $A$  must be equal to -1.

### 3.3 SUSY algebra

The algebra associated with the generators of a symmetry plays a fundamental role. Since we are working with spinor charges, we can expect anti-commutator relations. Here we will only refer the basic relations originally developed by Golfand and Likhtman [36]:

$$[Q, P^\mu] = \{Q_a, Q_b\} = \{Q_a^\dagger, Q_b^\dagger\} = 0 \quad (3.14)$$

$$\{Q_a, Q_b^\dagger\} = (\bar{\sigma}^\mu)_{ab} P_\mu \quad (3.15)$$

$Q_{a,b}$  is the SUSY generators while  $P_\mu$  is the momentum generator of space-time translations.

If  $Q_a$  is to be a real symmetry operator, it has to commute with the Hamiltonian, and this must also be true for the anti-commutator.

$$[Q_a, H] = [\{Q_a, Q_b\}, H] = 0 \quad (3.16)$$

SUSY is mainly thought of as a symmetry between fermions and bosons. While the author of [35] points out also this result (Eq. 3.16) as a remarkable feature of SUSY. It shows that the SUSY generators are connected to the momentum operator, which in turn is the generator of space-time displacement. SUSY seems therefore to be a kind of extension of space-time.

The development of what later became SUSY started in the 1960's when scientists were trying to combine internal symmetries with external Lorentz symmetries as a step towards a unified theory. But in 1967 it was shown by Coleman and Mandula that non-trivial

combination of internal and external symmetries (associated with the Poincaré algebra) could not be achieved using only bosonic charges [37]. SUSY algebra was then introduced as an extended Poincaré algebra with fermionic charges. As we will see later, unification and SUSY are closely related, and one could actually say that the Grand Unified Theory (GUT) is necessarily supersymmetric. The consequences of SUSY which will be presented later are a strong support for this statement.

### 3.4 Towards the Wess-Zumino model

In order to obtain an expression for the change of the fields one can start with calculating the difference between two subsequent variations  $(\delta_\eta\delta_\xi - \delta_\xi\delta_\eta)$  applied to  $\phi$ .

$$(\delta_\eta\delta_\xi - \delta_\xi\delta_\eta)\phi = (\xi^T C \sigma^\mu C \eta^* - \eta^T C \sigma^\mu C \xi^*) i \partial_\mu \phi \quad (3.17)$$

where the charge conjugation operator  $C$  is

$$C = -i\sigma_2 \quad (3.18)$$

One would like to obtain a consistent result when following a similar procedure for  $\chi$ . However it appears that the result is different:

$$(\delta_\eta\delta_\xi - \delta_\xi\delta_\eta)\chi_a = (\xi^T C \sigma^\mu C \eta^* - \eta^T C \sigma^\mu C \xi^*) i \partial_\mu \chi_a + i\xi_a (\eta^\dagger \bar{\sigma}^\mu \partial_\mu \chi) - i\eta_a (\xi^\dagger \bar{\sigma}^\mu \partial_\mu \chi) \quad (3.19)$$

Comparing to the result for  $\phi$  in Eq. 3.17, the problem lies in the additional two last terms. The equation of motion  $\bar{\sigma}^\mu \partial_\mu \chi = 0$  ensures that they vanish for massless, 'on-shell' fields. But we are also interested in 'off-shell' fields which correspond to internal lines in Feynman graphs. This is due to our choice of fields and the resulting mismatch of number of degrees of freedom in  $\phi$  and  $\chi$ . The easiest way out is to introduce a scalar field  $F$  to compensate for the difference.  $F$  is a so-called auxiliary field and we will have to use this maneuver once more before arriving at the complete SUSY Lagrangian. With the additional  $F$ -term the Lagrangian is now

$$\mathcal{L} = \partial_\mu \phi^\dagger \partial_\mu \phi + \chi^\dagger i \bar{\sigma}^\mu \partial_\mu \chi + F^\dagger F. \quad (3.20)$$

The  $F$ -field transforms like a  $\phi$ -type field:

$$\delta_\xi F = -i\xi^\dagger \bar{\sigma}^\mu \partial_\mu \chi \quad \delta_\xi F^\dagger = i\partial_\mu \chi^\dagger \bar{\sigma}^\mu \xi. \quad (3.21)$$

And the change in  $\chi$  can be written in its final form:

$$\delta_\xi \chi_a = -i\sigma^\mu (i\sigma_2 \xi^*) \partial_\mu \phi + \xi_a F \quad \delta_\xi \chi_a^\dagger = -i\partial_\mu \phi^\dagger \xi^T i\sigma_2 i\sigma^\mu + \xi_a^\dagger F^\dagger \quad (3.22)$$

Calculating now the  $\delta_\eta\delta_\xi\chi$ -terms in Eq. 3.19, the contributions from the  $F$ -terms in Eq. 3.22, which are proportional to the terms in Eq. 3.21, will cancel out the last two terms in Eq. 3.19. Adding the  $F$ -field and modifying the change in  $\chi$  makes the  $(\delta_\eta\delta_\xi - \delta_\xi\delta_\eta)\chi$  of the same form as the expression for  $\phi$ . Although this summary is too brief to justify all the technical details, its main purpose is to demonstrate why the auxiliary field  $F$  is needed.

## Interaction Lagrangian

Before proceeding to the interaction Lagrangian it is useful to introduce a more efficient notation:

$$\delta\phi_\xi = \xi^T(-i\sigma_2)\chi = \xi \cdot \chi \quad (3.23)$$

The Lagrangian describing a free left-handed supermultiplet with a massless spin-0 field  $\phi$  is given by Eq. 3.20. We now formulate interaction terms such that they preserve the invariance under SUSY transformation. In addition, they have to be renormalizable. The most general form is:

$$\mathcal{L}_{int} = W_i(\phi, \phi^\dagger)F_i - \frac{1}{2}W_{ij}(\phi, \phi^\dagger)\chi_i \cdot \chi_j + \text{h.c.} \quad (3.24)$$

where  $W_i$  and  $W_{ij}$  are defined below. We now have to ensure that  $\mathcal{L}_{int}$  is invariant under SUSY transformations. This requirement imposes several constraints and helps us to find the form of  $W$ . When applying the SUSY transformation to  $\mathcal{L}_{int}$  one obtains a  $\delta_\xi \mathcal{L}_{int}$  which contains several terms which will not cancel unless we define  $W_i$  and  $W_{ij}$  in an appropriate way. Thus we obtain:

$$W_i = \frac{\partial W}{\partial \phi_i} \quad W_{ij} = \frac{\partial^2 W}{\partial \phi_i \partial \phi_j} \quad (3.25)$$

$$W = \frac{1}{2}M_{ij}\phi_i\phi_j + \frac{1}{6}y_{ijk}\phi_i\phi_j\phi_k \quad (3.26)$$

$W$  is known as the superpotential and the term  $M_{ij}$  can be interpreted as mass, while  $y_{ijk}$  are the Yukawa couplings.

The Euler-Lagrange (E-L) equations for the fields  $F$  and  $F^\dagger$  which do not have a kinematic term give us the relations:

$$F_i = -W_i^\dagger \quad F_i^\dagger = -W_i \quad (3.27)$$

The Wess-Zumino Lagrangian containing both the free and interaction parts has then the form:

$$\mathcal{L}_{WZ} = \partial_\mu \phi_i^\dagger \partial^\mu \phi_i + \chi_i^\dagger i \bar{\sigma}^\mu \partial_\mu \chi_i + F_i^\dagger F_i - |W_i|^2 - \frac{1}{2}\{W_{ij}\chi_i \cdot \chi_j + \text{h.c.}\} \quad (3.28)$$

where the index  $i, j$  runs over internal degrees of freedom like flavor.

The term  $|W|^2 = |F|^2$  can be recognized as a part of the scalar potential  $V(\phi, \phi^\dagger)$  and it is therefore called the “F-component”. If we write out the products we will find a quadratic self-coupling with a  $(\text{mass})^2$  term and cubic and quartic terms with different combinations of mass and Yukawa couplings.

Applying the E-L equations for the different fields in  $\mathcal{L}_{WZ}$  we see that the quadratic parts describe free spin-0 and spin- $\frac{1}{2}$  fields with degenerate masses. By introducing interaction and thus the superpotential, we also introduced mass.

## 3.5 SUSY as a gauge theory

So far we have only looked at the matter fields while SUSY must also provide partners for the SM gauge bosons, the so called gauginos. In SM we have the massless U(1) spin-1



photon field  $A_\mu(x)$ . Now we have to add its fermionic partner, photino ( $\tilde{\gamma}$ ) which must be a spin- $\frac{1}{2}$  left-handed field  $\lambda$ . The corresponding Lagrangian is

$$\mathcal{L}_{\gamma\lambda} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\lambda^\dagger\bar{\sigma}^\mu\partial_\mu\lambda \quad (3.29)$$

where  $F^{\mu\nu}$  is the usual field strength of the SM.

We now look at the field transformations.  $A_\mu$  is a bosonic field and the following is proposed:

$$\delta_\xi A^\mu = \xi^\dagger\bar{\sigma}^\mu\lambda + \lambda^\dagger\bar{\sigma}^\mu\xi \quad (3.30)$$

While for  $\lambda$ :

$$\delta_\xi\lambda \sim \xi F^{\mu\nu} \quad (3.31)$$

where  $F^{\mu\nu}$  is used rather than  $A_\mu$  since it is gauge invariant. Again, we require that  $\mathcal{L}_{\gamma\lambda}$  is invariant under SUSY transformations which gives them the form

$$\delta_\xi\lambda = \frac{1}{2}i\sigma^\mu\bar{\sigma}^\nu\xi F^{\mu\nu} \quad \delta_\xi\lambda^\dagger = -\frac{1}{2}i\xi^\dagger\bar{\sigma}^\nu\sigma^\mu F^{\mu\nu} \quad (3.32)$$

However, the calculations of  $(\delta_\eta\delta_\xi - \delta_\xi\delta_\eta)$  applied to  $A_\mu$  and  $\lambda$  meet the same problems as outlined in Section 3.4 for  $\chi$ . There is therefore a need for one more auxiliary field, analogous to the  $F$ -field. It is commonly known as the  $D$ -field and it is a non-propagating real scalar field with one degree of freedom. The Lagrangian now gets an extra term:

$$\mathcal{L}_D = \frac{1}{2}D^2. \quad (3.33)$$

The transformation of  $D$  follows the same pattern as  $F$ , and  $\delta_\xi\lambda$  acquires an additional term which ensures the invariance of the Lagrangian.

$$\delta_\xi D = -i(\xi^\dagger\bar{\sigma}^\mu\partial_\mu\lambda - (\partial_\mu\lambda)^\dagger\bar{\sigma}^\mu\xi) \quad (3.34)$$

$$\delta_\xi\lambda = \frac{1}{2}i\sigma^\mu\bar{\sigma}^\nu\xi F^{\mu\nu} + \xi D \quad \delta_\xi\lambda^\dagger = -\frac{1}{2}i\xi^\dagger\bar{\sigma}^\nu\sigma^\mu F^{\mu\nu} + \xi^\dagger D \quad (3.35)$$

Now that we introduced the auxiliary field, we need to find a supersymmetric version of the SU(2) non-abelian gauge theory. A general field strength can be written as:

$$\mathbf{F}_{\mu\nu} = \partial_\mu\mathbf{W}_\nu - \partial_\nu\mathbf{W}_\mu - g\mathbf{W}_\mu \times \mathbf{W}_\nu \quad (3.36)$$

where  $g$  is the generalized charge. The vector field for SU(2) is  $\mathbf{W} = (W_\mu^1, W_\mu^2, W_\mu^3)$ , while it would contain the 8 gluon fields in the case of SU(3). After an infinitesimal local gauge transformation it takes the form

$$W_\mu'^a(x) = W_\mu^a(x) - \partial_\mu\alpha^a(x) - g\epsilon_{abc}\alpha^b(x)W_\mu^c(x) \quad (3.37)$$

where  $\epsilon_{abc}$  is the structure constant, while  $\alpha^a(x)$  is the infinitesimal local transformation. This expression is analogous to Eq. 2.25. Details on how  $\lambda^a$  and  $F_{\mu\nu}^a$  transform can be found in [35]. A generalized version of the gauge Lagrangian for vector fields is:

$$\mathcal{L}_{W\lambda} = -\frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu} + i\lambda^{a\dagger}\bar{\sigma}^\mu\partial_\mu\lambda^a \quad (3.38)$$

Investigating how this Lagrangian behaves under gauge transformations one realizes that the last term is not invariant. The standard remedy is to replace the derivative with the covariant derivative

$$\partial_\mu \rightarrow D_\mu \equiv \partial_\mu + ig\mathbf{T}\mathbf{W}_\mu. \quad (3.39)$$

The  $\mathbf{T}$  matrices are the representations of the generators of the groups (Pauli matrices in case of SU(2) and Gell-Mann  $\lambda^a$  matrices for SU(3)). After this substitution we require the resulting gauge invariant  $\mathcal{L}_{gauge}$  to also be invariant under SUSY transformations. In analogy to the U(1) case, we have to add an auxiliary field, but in case of SU(2) one needs a triplet of  $D$ 's,  $D^a$ , to compensate for the off-shell degree of freedom for each  $W_\mu^a$ . The proposal for a SUSY- and gauge-invariant Lagrangian is

$$\mathcal{L}_{gauge} = -\frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu} + i\lambda^{a\dagger}\bar{\sigma}^\mu(D_\mu\lambda)^a + \frac{1}{2}D^a D^a \quad (3.40)$$

$D$  is a real scalar field,  $\lambda$  is a fermionic field and  $W_\mu^a$  vector fields. The SUSY transformations of these fields are as follows:

$$\begin{aligned} \delta_\xi W^{\mu a} &= \xi^\dagger \bar{\sigma}^\mu \lambda^a + \lambda^{a\dagger} \bar{\sigma}^\mu \xi \\ \delta_\xi \lambda^a &= \frac{1}{2} i \sigma^\mu \bar{\sigma}^\nu \xi F_{\mu\nu}^a + \xi D^a \\ \delta_\xi D^a &= -i(\xi^\dagger \bar{\sigma}^\mu (D_\mu \lambda)^a - (D_\mu \lambda)^{a\dagger} \bar{\sigma}^\mu \xi) \end{aligned} \quad (3.41)$$

The gauge Lagrangian in Eq. 3.40 expressed in terms of covariant derivatives is now invariant both under gauge transformations and SUSY transformations (Eq. 3.41).

### 3.5.1 Towards the Minimal Supersymmetric Standard Model - MSSM

So far we have been treating the Lagrangian for matter and gauge fields separately, while a realistic theory must include interactions between these two kinds of fields. For U(1) we introduce the covariant derivative

$$D_\mu = \partial_\mu + iqA_\mu \quad (3.42)$$

where  $q$  is the U(1) coupling constant (the electric charge  $e$ ) and obtain the following Lagrangian

$$\mathcal{L} = (D_\mu \phi)^\dagger (D^\mu \phi) + i\chi^\dagger \bar{\sigma}^\mu D_\mu \chi + F^\dagger F - \frac{1}{4}F_{\mu\nu} F^{\mu\nu} + i\lambda^\dagger \bar{\sigma}^\mu \partial_\mu \lambda + \frac{1}{2}D^2. \quad (3.43)$$

The next step is to add a term which describes the interaction between the matter fields  $\phi$  and  $\chi$  and the fields  $\lambda$  and  $D$ . It is defined by requiring that it should be Lorentz invariant, renormalizable and gauge invariant. Guided by dimensional arguments (dimensions of mass) one can write a general interaction term as

$$Aq[(\phi^\dagger \chi) \cdot \lambda + \lambda^\dagger \cdot (\chi^\dagger \phi)] + Bq\phi^\dagger \phi D \quad (3.44)$$

where the coefficients  $A$  and  $B$  are left to be found [35] by requiring that the combined Lagrangian in Eq. 3.43 plus the interaction term Eq. 3.44 is still invariant under SUSY transformations. It appears that the  $\xi$  applied in the variation related to the SUSY

transformations is not necessarily the same as earlier. This leads to the introduction of a scaling factor in the transformations. The requirement that the interaction (Eq. 3.44) should be SUSY-invariant provides the unknown factors  $A = -\sqrt{2}$  and  $B = -1$  and in addition modifies the variation of the  $F$  field. Finally we obtain the SUSY-invariant Lagrangian which combines the chiral fields and the U(1) super-multiplet

$$\begin{aligned} \mathcal{L} = & (D_\mu \phi)^\dagger (D^\mu \phi) + i\chi^\dagger \bar{\sigma}^\mu D_\mu \chi + F^\dagger F - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} + i\lambda^\dagger \bar{\sigma}^\mu \partial_\mu \lambda + \frac{1}{2} D^2 \\ & - \sqrt{2}q[(\phi^\dagger \chi) \cdot \lambda + \lambda^\dagger \cdot (\chi^\dagger \phi)] - q\phi^\dagger \phi D \end{aligned} \quad (3.45)$$

### SU(2) the non-abelian case

The starting point for the non-abelian case is the Wess-Zumino Lagrangian with the superpotential  $W$

$$\mathcal{L}_{WZ} = \partial_\mu \phi_i^\dagger \partial^\mu \phi_i + \chi_i^\dagger i\bar{\sigma}^\mu \partial_\mu \chi_i + F_i^\dagger F_i + \left[ \frac{\partial W}{\partial \phi_i} F_i - \frac{1}{2} \frac{\partial^2 W}{\partial \phi_i \partial \phi_j} \chi_i \cdot \chi_j + \text{h.c.} \right] \quad (3.46)$$

in which we replace the normal derivatives by the covariant derivatives

$$\begin{aligned} D_\mu \phi_i &= \partial_\mu \phi_i + igA_\mu^a (T^a \phi)_i \\ D_\mu \chi_i &= \partial_\mu \chi_i + igA_\mu^a (T^a \chi)_i \end{aligned} \quad (3.47)$$

where  $T^a$  are the generators of the group,  $g$  is the gauge coupling constant and  $A_\mu^a$  is the gauge field. The procedure for finding the interaction terms is the same as for U(1) and the result is

$$\begin{aligned} \mathcal{L}_{\text{gauge+chiral}} &= \mathcal{L}_{\text{gauge}} + \mathcal{L}_{WZ} \\ &- \sqrt{2}g[(\phi_i^\dagger T^a \chi_i) \cdot \lambda^a + \lambda^{a\dagger} \cdot (\chi_i^\dagger T^a \phi_i)] - g(\phi_i^\dagger T^a \phi_i) D^a \end{aligned} \quad (3.48)$$

where  $\mathcal{L}_{WZ}$  is (3.46) and  $\mathcal{L}_{\text{gauge}}$  is (3.40). And the equation of motion for  $D^a$  is

$$D^a = g \sum_i \phi_i^\dagger T^a \phi_i. \quad (3.49)$$

This can be used to eliminate the auxiliary field  $D$  and the scalar potential  $V(\phi, \phi^\dagger)$  which enters the definition of the Lagrangian ( $L = T - V$ ) is

$$V(\phi, \phi^\dagger) = \sum_i |F_i|^2 + \sum_\alpha \frac{1}{2} (D^\alpha)^2 = \sum_i |W_i|^2 + \frac{1}{2} \sum_G \sum_a \sum_{i,j} g_G^2 (\phi_i^\dagger T_G^a \phi_i) (\phi_j^\dagger T_G^a \phi_j) \quad (3.50)$$

$W_i$  is given by Eq. 3.25 and 3.26 and contains the fermion masses  $M_{ij}$  and Yukawa couplings. It is often referred to as the  $F$ -term, while the second term, the  $D$ -term, is determined by the gauge interactions.

## 3.6 The Minimal Supersymmetric Standard Model MSSM

SUSY does not introduce additional interactions to the four familiar ones and the coupling constants are the same as for the SM. However, the particle spectrum is doubled. All the

fermions acquire supersymmetric partners, sleptons and squarks, where the left-handed fields  $\chi$  are  $SU(2)_L$ -doublets, while the right-handed  $\psi \sim i\sigma_2 \chi^{\dagger TC}$  are  $SU(2)_L$ -singlets. The building blocks are now the chiral super-fields and we use the 4-component Dirac spinors applying the charge conjugation operator  $C_0 = -i\gamma_2$  in order to describe the right-handed fields in terms of left-handed ones.

The gauge fields of the unbroken MSSM are summarized in table 3.1.

| Super-multiplets                        | Bosonic Fields spin-0  | Fermionic Fields spin-1/2  |
|---|--|--|
| Q, 3 families<br>$\bar{u}$<br>$\bar{d}$ | Squarks $\tilde{Q} = (\tilde{u}_L, \tilde{d}_L)$<br>Squarks $\tilde{\bar{u}}_L(\tilde{u}_R)$<br>Squarks $\tilde{\bar{d}}_L, (\tilde{d}_R)$ | Quarks $Q = (u_L, d_L)$<br>Quarks $\bar{u}_L \sim (u_R)^C$<br>Quarks $\bar{d}_L \sim (d_R)^C$                      |
| L, 3 families<br>$\bar{e}$              | Sleptons $\tilde{L} = (\tilde{\nu}_{eL}, \tilde{e}_L)$<br>Sleptons $\tilde{\bar{e}}_L(\tilde{e}_R)$  | Leptons $L = (\nu_{eL}, e_L)$<br>Leptons $\bar{e}_L \sim (e_R)^C$  |
| g<br>W<br>B                             | Gluon $g$<br>W-boson $W^\pm$<br>B-boson $B$  | Gluino $\tilde{g}$<br>Wino $\tilde{W}^\pm$<br>Bino $\tilde{B}$   |
| $H_u$<br>$H_d$                          | Higgs $H_u = (H_u^+, H_u^0)$<br>Higgs $H_d = (H_d^0, H_d^-)$   | Higgsino $\tilde{H}_u = (\tilde{H}_u^+, \tilde{H}_u^0)$<br>Higgsino $\tilde{H}_d = (\tilde{H}_d^0, \tilde{H}_d^-)$ |

Table 3.1: List of fields in the MSSM.

The physical content of particles and gauge fields in the theory suggests an appropriate form of the superpotential [35]

$$W = y_u^{ij} \bar{u}_i Q_j \cdot H_u - y_d^{ij} \bar{d}_i Q_j \cdot H_d - y_e^{ij} \bar{e}_i L_j \cdot H_d + \mu H_u \cdot H_d \quad (3.51)$$

The Yukawa couplings are the same as in SM and  $W$  provides the SM matter particles with their familiar masses. The “ $\cdot$ ” means  $SU(2)$  invariant coupling of two doublets analogous to the Lorentz invariant spinor product. Color indices are suppressed.

### 3.6.1 Spontaneous SUSY breaking

SUSY is held to be a broken symmetry since we have not observed any supersymmetric particles yet and therefore assume they must be heavier than the SM particles. The breaking is established by adding all possible terms which break SUSY without reintroducing quadratic divergences. These are [34]<sup>2</sup>:

$$-\mathcal{L}_{gaugino} = \frac{1}{2} [M_1 \tilde{B} \tilde{B} + M_2 \sum_{a=1}^3 \tilde{W}^a \tilde{W}_a + M_3 \sum_{a=1}^8 \tilde{G}^a \tilde{G}_a + \text{h.c.}] \quad (3.52)$$

$$-\mathcal{L}_{fermions} = + \sum_{i=\text{gen}} m_{\tilde{Q}_i}^2 \tilde{Q}_i^\dagger \tilde{Q}_i + m_{\tilde{L}_i}^2 \tilde{L}_i^\dagger \tilde{L}_i + m_{\tilde{u}_i}^2 |\tilde{u}_{R_i}|^2 + m_{\tilde{d}_i}^2 |\tilde{d}_{R_i}|^2 + m_{\tilde{l}_i}^2 |\tilde{l}_{R_i}|^2 \quad (3.53)$$

$$-\mathcal{L}_{Higgs} = m_{H_u}^2 H_u^\dagger H_u + m_{H_d}^2 H_d^\dagger H_d + b(H_u H_d + \text{h.c.}) \quad (3.54)$$

$$-\mathcal{L}_{tril.} = \sum_{i,j=\text{gen.}} [A_{ij}^u y_{ij}^u \tilde{u}_{R_i} H_u \tilde{Q}_j + A_{ij}^d y_{ij}^d \tilde{d}_{R_i} H_d \tilde{Q}_j + A_{ij}^l y_{ij}^l \tilde{l}_{R_i} H_d \tilde{L}_j + \text{h.c.}] \quad (3.55)$$

<sup>2</sup>The parameter  $b$  in Eq. 3.54 is in [34] called  $B\mu$ .

The SUSY breaking happens in a so-called hidden sector and is mediated to the visible MSSM sector. SUSY cannot be broken in a satisfactory way by a global symmetry transformation. While, on the other hand, applying local transformations leads to a spontaneous symmetry breaking mechanism which necessarily also involves gravity. The local character of the transformation introduces a spin-2 graviton and spin-3/2 gravitino. Analogous to the SM Higgs mechanism, one introduces a massless Goldstone boson. In this case it is the massless gravitino that 'eats' a Goldstino (1/2-spin Goldstone fermion) and acquires mass, while the graviton remains massless also after the breaking of SUSY.

One of the most common models is the so-called minimal supergravity (mSUGRA) in which the SUSY-breaking is mediated to the visible sector by the flavor-blind gravitational interaction. There are also other SUSY-breaking mediation mechanisms as for example Gauge Mediated SUSY-breaking (GMSB) and Anomaly Mediated SUSY-breaking (AMSB).

The unconstrained MSSM introduces 105 free parameters in addition to the 19 involved in the SM. The number of parameters can be reduced if the SUSY breaking parameters obey a set of universal boundary conditions at the GUT scale. This is the topic of the next section.

### 3.6.2 Gauge coupling unification

As mentioned earlier, SUSY is closely related to unification. If the interactions unify at a certain energy scale, we expect the running coupling constants that set their strength to meet there. The constants are defined as <sup>3</sup>

$$\alpha_3 = \frac{g_s^2}{4\pi} \quad \alpha_2 = \frac{g^2}{4\pi} \quad \alpha_1 = \frac{5g'^2}{3 \cdot 4\pi} = \frac{g_1^2}{4\pi}, \quad g \tan \theta_W = g' = \sqrt{\frac{3}{5}} g_1. \quad (3.56)$$

The last relation shows that coupling constants are closely related to the weak mixing angle, which is measured to be  $\sin^2 \theta_W = 0.22331 \pm 0.00062$  [24] at the  $Z$  peak. In order to see how unification works, if the couplings indeed meet, the measured value of  $\sin^2 \theta_W$  is used as input to the renormalization Group Equations (RGEs). Unfortunately, within the SM the crossing points are rather scattered as shown in the upper plot of Figure 3.2. A unification would only be possible if the value of  $\sin^2 \theta_W$  was  $0.214 \pm 0.004$  at the EW scale [33]. Minimal supersymmetric GUT, on the other hand, predicts unification with  $\sin^2 \theta_W \simeq 0.232$ , which is in good agreement with the measured value, and the unification is more precise as shown in lower plot of Figure 3.2. These results represent some of the strongest arguments in favour of supersymmetry.

We now take a closer look at the running of the constants as described by the RGEs in order to see how the additional supersymmetric particles change their evolution. To one loop order they have the form:

$$\frac{d}{d \ln Q}(\alpha_i^{-1}) = \frac{b_i}{2\pi} \quad (3.57)$$

After integrating we arrive to:

$$\alpha_i^{-1}(Q) = \alpha_i^{-1}(Q_0) + \frac{b_i}{2\pi} \ln(Q/Q_0). \quad (3.58)$$

---

<sup>3</sup> $\alpha_3$  is for  $SU(3)_C$ ,  $\alpha_2$  -  $SU(2)_L$  and  $\alpha_1$  for  $U(1)$

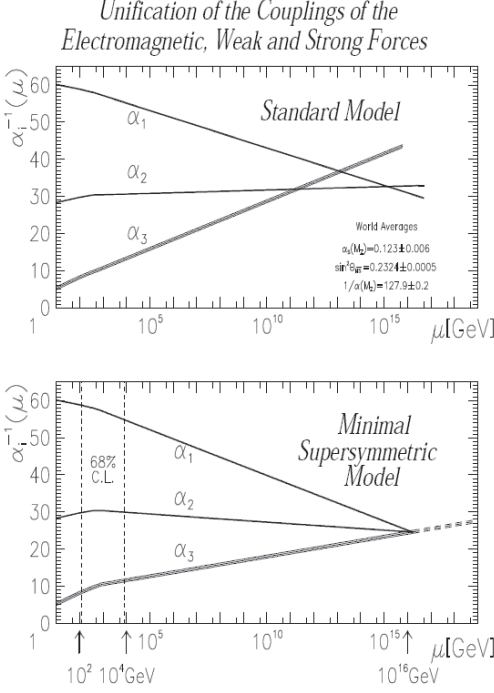


Figure 3.2: Unification of the running coupling constants at the GUT scale. [38]

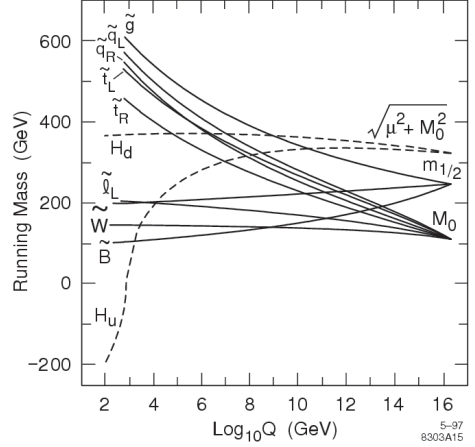


Figure 3.3: Evolution of the soft SUSY breaking masses. The assumption is that the masses are unified at GUT scale and evolve by RGE down to EW scale.  $m_{1/2}$  is the common scalar mass,  $M_0$  the common gaugino mass and  $\mu^2$  the Higgsino mass parameter [39].

$Q$  is the running energy scale while  $Q_0$  is the initial value at which  $\alpha_i$  is measured. The improved convergence point which we see in the lower plot of Figure 3.2 is due to the change of slope of the curves caused by the factor  $b_i$  which is defined as

$$b_i = \frac{11}{3}N - \frac{1}{3}n_f - \frac{1}{6}n_s. \quad (3.59)$$

where  $i$  denotes the gauge group,  $n_f$  is the number of fermions and  $n_s$  the number of scalars available at a given energy scale. With additional supersymmetric particles at around 1 TeV the slope changes and the unification is predicted at an energy scale  $M_{GUT} \sim 2 \cdot 10^{16}$  GeV.

One can also apply a similar RGE evolution of the masses. As an example we look at the gaugino masses  $M_{1,2,3}$ , defining  $t \equiv \ln Q$ :

$$\frac{dM_i}{dt} = -\frac{b_i}{2\pi}\alpha_i M_i \quad (3.60)$$

Using Eq. 3.58 we obtain

$$\frac{d}{dt}(M_i/\alpha_i) = 0 \quad (3.61)$$

which means that  $M_i/\alpha_i$  are scale-independent and we obtain the following hierarchy at  $m_Z$ :

$$M_1(m_Z) = \frac{\alpha_1(m_Z)}{\alpha_2(m_Z)}M_2(m_Z) \simeq 0.5M_2(m_Z), \quad M_3(m_Z) = \frac{\alpha_3(m_Z)}{\alpha_2(m_Z)}M_2(m_Z) \simeq 3.5M_2(m_Z) \quad (3.62)$$

This leads to the following hierarchy and electroweak scale

$$M_3(m_Z) : M_2(m_Z) : M_1(m_Z) \sim 7 : 2 : 1. \quad (3.63)$$

Following the RGE development towards unification of masses, as shown in Figure 3.3, leads to a constrained model like mSUGRA with common masses at the unification scale:

- Unification of gaugino masses:  
 $M_i(M_{GUT}) = M_2(M_{GUT}) = M_3(M_{GUT}) \equiv m_{1/2}$
- Unification of scalar masses:  
 $m_{\tilde{Q}}(M_{GUT}) = m_{\tilde{u}_R}(M_{GUT}) = m_{\tilde{d}_R}(M_{GUT}) =$   
 $m_{\tilde{L}}(M_{GUT}) = m_{\tilde{t}_R}(M_{GUT}) = m_{H_u}(M_{GUT}) = m_{H_d}(M_{GUT}) \equiv m_0$

The unknown parameters associated with the Higgs sector can be expressed in terms of other variables leading to the reduced mSUGRA parameter space:  $\tan \beta$ ,  $m_{1/2}$ ,  $m_0$ ,  $A_0$  and  $\text{sign}(\mu)$ . The angle  $\beta$  will be defined in a moment.

Although the reduced parameter scheme of mSUGRA is a useful simplification of the complex phenomenology of SUSY, it represents at the same time a limitation. Chapter 10 will therefore describe a more general model MSSM24 which suffers less from the strict mass hierarchy imposed by mSUGRA.

## 3.7 SUSY particle masses

What we have been working with so far are the gauge eigenstates, while the physical particles which may be observed in the experiments are mass eigenstates constructed through a mixing of the gauge eigenstates.

The mass mixing matrices can be found in many publications, for example [34]. They are included here in order to show their structure, list the expressions for SUSY particle masses as well as the practical simplified formulae related to the model parameters. The mass hierarchy and the corresponding gauge field composition play an important role in SUSY phenomenology as they define which production processes and decays are favoured.

In the following we use the short-hand notation:  $s_\beta \equiv \sin \beta$ ,  $c_\beta \equiv \cos \beta$ ,  $s_W \equiv \sin \theta_W$  and  $c_W \equiv \cos \theta_W$

### The Higgs sector

MSSM requires two complex Higgs doublets:

$$H_d = \begin{pmatrix} H_d^0 \\ H_d^- \end{pmatrix} \quad Y_{H_d} = -1 \quad H_u = \begin{pmatrix} H_u^+ \\ H_u^0 \end{pmatrix} \quad Y_{H_u} = +1 \quad (3.64)$$

which give 8 real degrees of freedom. Three of them give mass to the SM gauge bosons, while the remaining 5 lead to physical Higgs particles which one hopes to discover. The theory contains two neutral, CP-even Higgs bosons  $h$  and  $H$ , two charged  $H^\pm$  and the CP-odd, neutral  $A$ . Similar to SM, but in a more complicated way, neutral components acquire vacuum expectation values:

$$\langle H_u^0 \rangle = v_u \quad \langle H_d^0 \rangle = v_d \quad (3.65)$$

These add up to give the electroweak scale

$$\sqrt{v_u^2 + v_d^2} = \left( \frac{2m_Z^2}{g^2 + g'^2} \right)^{1/2} = 174 \text{ GeV} \quad (3.66)$$

and we define the  $\beta$  parameter introduced above as

$$\tan \beta \equiv \frac{v_u}{v_d} \quad (3.67)$$

which is a convenient parameter in further calculations [40].

We are now interested in physical particles which are mass eigenstates. A typical mass term in the Lagrangian is

$$\mathcal{L} \propto -\psi^T \mathbf{M} \psi + c.c. \quad (3.68)$$

where  $\psi$  is the gauge eigenstate basis vector. To obtain the masses one has to diagonalize the  $\mathbf{M}$  matrix by a unitary matrix  $\mathbf{N}$  which is related to the mass eigenstates through  $\phi_i = N_{ij} \psi_j$ . This general procedure will be used to find the masses of the supersymmetric particles.

The masses in the Higgs sector are described by several parameters, but as they are mutually related, one can reduce the number to two free parameters. The relations impose also a strongly constrained mass hierarchy. A common choice of free parameters is  $m_A$  and  $\tan \beta$ . For the sake of completeness, the list below shows the formulae for all Higgs masses

$$m_A^2 = 2b/s_{2\beta} \quad (3.69)$$

$$m_{H^\pm}^2 = m_A^2 + m_W^2 \quad (3.70)$$

$$m_{h,H^0}^2 = \frac{1}{2} \left[ m_A^2 + m_Z^2 \mp \sqrt{(m_A^2 + m_Z^2)^2 - 4m_A^2 m_Z^2 c_{2\beta}^2} \right] \quad (3.71)$$

where  $b$  is a free parameter in the Higgs scalar potential [40]. In Eq. 3.71, minus is for  $h$ , plus for  $H^0$ . This is an important result as it sets the upper bound on the mass of the lightest Higgs. Including quantum correction where the top-stop one-loop correction is typically the largest contributor [41–43], leads to the constraint:

$$m_h \leq m_Z \cdot |c_{2\beta}| \lesssim 130 \text{ GeV}. \quad (3.72)$$

This prediction of a relatively light Higgs is in good agreement with other calculations and limits, for example [24, 44], and provides yet another argument in favour of SUSY.

### Charginos

Charginos are mixtures of the charged winos and higgsinos which form the gauge eigenstate basis. The chargino mass matrix is given by:

$$M_{\chi^\pm} = \begin{pmatrix} M_2 & \sqrt{2}m_W s_\beta \\ \sqrt{2}m_W c_\beta & \mu \end{pmatrix} \xrightarrow{\text{Diag.}} \begin{pmatrix} m_{\tilde{\chi}_1^\pm} & 0 \\ 0 & m_{\tilde{\chi}_2^\pm} \end{pmatrix}$$

which in the limit  $|\mu| \gg M_2, m_W$  leads to the reduced expressions:

$$m_{\tilde{\chi}_1^\pm} \simeq M_2 - \frac{m_W^2}{\mu^2} (M_2 + \mu s_{2\beta}) \quad m_{\tilde{\chi}_2^\pm} \simeq |\mu| + \frac{m_W^2}{\mu^2} \epsilon_\mu (M_2 s_{2\beta} + \mu) \quad (3.73)$$



where  $\epsilon_\mu \equiv \text{sign}\mu$ . From this simplified form we can see that when  $|\mu| \rightarrow \infty$ , the lightest chargino is mostly wino  $m_{\chi_1^\pm} \simeq M_2$  while the heavier one is mostly higgsino  $m_{\chi_2^\pm} \simeq |\mu|$ . If  $M_2 \gg |\mu|, m_Z$  the gauge composition is opposite.

### Neutralinos

The basis in the neutralino sector consists of the neutral fields  $(\tilde{B}, \tilde{W}^0, \tilde{H}_1^0, \tilde{H}_2^0)$ . The mixing matrix is

$$M_{\chi^0} = \begin{pmatrix} M_1 & 0 & -m_Z s_W c_\beta & m_Z s_W s_\beta \\ 0 & M_2 & m_Z c_W c_\beta & -m_Z c_W s_\beta \\ -m_Z s_W c_\beta & m_Z c_W c_\beta & 0 & -\mu \\ m_Z s_W s_\beta & -m_Z c_W s_\beta & -\mu & 0 \end{pmatrix}$$

and if  $|\mu|$  is large and  $|\mu| \gg M_{1,2} \gg M_Z$  the expressions for the neutralino masses simplify to [34]:

$$m_{\chi_1^0} \simeq M_1 - \frac{m_Z^2}{\mu^2}(M_1 + \mu s_{2\beta})s_W^2, \quad m_{\chi_2^0} \simeq M_2 - \frac{m_Z^2}{\mu^2}(M_2 + \mu s_{2\beta})c_W^2 \quad (3.74)$$

$$m_{\chi_{3,4}^0} \simeq |\mu| + \frac{1}{2} \frac{m_Z^2}{\mu^2} \epsilon_\mu (1 \mp s_{2\beta}) (\mu \pm M_2 s_W^2 \mp M_1 c_W^2) \quad (3.75)$$

When  $|\mu|$  is large, the two lightest neutralinos are mostly gaugino like (bino and wino) with masses  $\chi_1^0 \approx M_1$ ,  $\chi_2^0 \approx M_2$ . The two heavier neutralinos are then mostly Higgsino like with mass  $m_{\chi_{3,4}^0} \simeq |\mu|$ . When on the other hand  $|\mu|$  is small compared to  $m_Z$ , the gauge composition of the neutralinos is reversed.

### Gluino

The gluino mass at tree-level is [34]:

$$m_{\tilde{g}} = M_3 \quad (3.76)$$

Recalling the mass hierarchy of mSUGRA in Eq. 3.63, the gluino is expected to be much heavier than the lighter charginos and neutralinos.

### Sfermions

The left- and right-handed components interact differently leading to different masses. In general, we see from the mass formulae after diagonalization of the mass matrix, the right-handed sfermion is always lighter. The off-diagonal terms are proportional to the mass and the mixing effect is strongest for stop  $\tilde{t}$ . However, for higher values of  $\tan\beta$  the mass splitting can also be enhanced for  $\tilde{\tau}$  and  $\tilde{b}$ . Current experimental results tend to exclude low values of  $\tan\beta$ , so the mass splitting for  $\tilde{\tau}$  and  $\tilde{b}$  needs to be taken into account.

In constrained models like mSUGRA one assumes an universal scalar mass  $m_0$  and gaugino mass  $m_{1/2}$  which leads to simplified expressions. Performing the RGE evolution down to the electroweak scale leads to analytic approximations [45] of the masses for the first two families (using  $\alpha_{GUT} \simeq 0.041$ ) [34]:

$$m_{\tilde{q}_i}^2 \sim m_0^2 + 6m_{1/2}^2, \quad m_{\tilde{t}_L}^2 \sim m_0^2 + 0.52m_{1/2}^2, \quad m_{\tilde{e}_R}^2 \sim m_0^2 + 0.15m_{1/2}^2, \quad (3.77)$$

For  $\tilde{t}$  and  $\tilde{b}$  the mass splitting is non-negligible. Including the Yukawa coupling in the RGE evolution gives in the small  $\tan\beta$  regime

$$m_{\tilde{t}_L}^2 = m_{\tilde{b}_L}^2 \sim m_0^2 + 6m_{1/2}^2 - \frac{1}{3}X_t \quad m_{\tilde{t}_R}^2 = m_{\tilde{b}_R}^2 \sim m_0^2 + 6m_{1/2}^2 - \frac{2}{3}X_t \quad (3.78)$$

where  $X_t \sim 1.3m_0^2 + 3m_{1/2}^2$ .

### 3.8 R-parity

In order to ensure lepton and baryon number conservation in SUSY one introduces a new multiplicative symmetry: the R-parity, which is defined by

$$R = (-1)^{2s+3B+L} \quad (3.79)$$

Where  $s$  is spin,  $B$  baryon number and  $L$  lepton number.  $R = +1$  for SM particles and  $R = -1$  for supersymmetric partners. Conservation of R-parity has some important phenomenological consequences for SUSY:

- Sparticles are always produced in pairs.
- The decay of a sparticle includes an odd number of sparticles.
- The Lightest Supersymmetric Particle (LSP) is stable.

In many SUSY models the LSP is the neutralino, a weakly interacting particle which is considered to be a good dark matter candidate. If R-parity is conserved, there should be a great number of these left from Big Bang. The relic density  $\Omega_\chi h^2$  depends on the rate of processes like  $\chi_1^0 \chi_1^0 \rightarrow f\bar{f} \rightarrow \gamma\gamma$  which is again related to the couplings and masses, and thus to the  $m_0$  and  $m_{1/2}$  parameters and to the composition of the LSP. Therefore, a measurement of  $\Omega_\chi h^2$  can be used to set limits on the LSP mass and thus constrain the parameter space of the studied SUSY models.

This dependency is a good example of how closely particle physics and astrophysics are inter-dependent. However, it should be mentioned that there are also R-parity breaking SUSY models on the market, that lead to interesting phenomena like for example R-hadrons [46, 47].

### 3.9 SUSY searches, limits and mSUGRA constraints

The LEP experiments performed detailed SUSY searches mainly within the mSUGRA framework. In the absence of a signal the studies lead to lower limits on particle masses and model parameters. Detailed listing of results can be found in [48].

The parameter space of the various SUSY models have during the last years been significantly constrained by measurements sensitive to SUSY [49–51]. The main constraining observations are: the dark matter density which has been reported with increasing accuracy by the Wilkinson Microwave Anisotropy Probe (WMAP) [52], branching fraction of  $B \rightarrow s\gamma$  [53], anomalous magnetic moment of the muon ( $g_\mu - 2$ ) [54] and the LEP constraints, both on the Higgs masses and on SUSY. The low mass or so-called bulk-region is very much reduced and areas compatible with the  $\Omega_\chi h^2$  take the shape of narrow

stripes. Figure 3.4 (left) shows one example of a  $\chi^2/dof$  two dimensional scan of the mSUGRA model in the  $m_0 - m_{1/2}$  plane. The green/light gray regions correspond to the most favourable areas of the parameter space. One of the allowed regions is the high  $m_0$ -area corresponding to the so-called Focus point scenario. This study includes also limits from direct Weekly Interacting Massive Particles (WIMP) detection experiments CDMS [55] and GENIUS [56].

Both D0 and CDF experiments at the Tevatron search for SUSY [57] and one of the main channels is the trilepton final state. So far there has not been any observation of signal, however, both experiments are able to push some of the limits set by LEP. Figure 3.4 (right) shows a recent plot of expected and observed limit for the mSUGRA model ( $m_0 = 60$  GeV,  $\tan\beta = 3$  and  $A_0 = 0$ ). The value of  $m_{1/2}$  which sets the chargino mass has been varied. This has lead to an exclusion of chargino masses of less than 164 GeV at the 95% C.L. [58]. The D0 experiment reports on similar searches setting slightly lower but very similar limits [59]. They also study the  $\sigma(\chi_2^\pm\chi^0) \times BR(3l)$  as a function of  $\tan\beta$  for chargino mass of 130 GeV. The result excludes charginos of this mass up to  $\tan\beta = 9.6$ . The LEP limits have also been used to exclude gluinos with mass  $m_{\tilde{g}} \lesssim 550$  GeV [60].

It should be mentioned that it is difficult to compare the Tevatron results with the benchmark points studied in ATLAS. However, ATLAS has defined one low mass point which is very close to the Tevatron models.

Additional constraints may also come from astro-particle physics observations. If dark matter consists of neutralinos these may through annihilation processes lead to an excess of various particles which would be visible above the cosmic radiation from other sources.

Observing the galactic center, the EGRET telescope reported an excess of  $\gamma$ -ray flux which could be interpreted as a result of pair annihilation of WIMPs. Many analyses suggest these could be supersymmetric particles [61]. However, it has not been possible to discriminate between the possible sources of this excess and there are analyses showing that these are compatible with more ordinary cosmic radiation [62]. The much more powerful Fermi Gamma-ray Space Telescope (earlier called GLAST) [63] is currently collecting gamma-ray data and is expected to provide valuable insight and limits in the near future.

In addition there is an ongoing study of positron and antiproton spectrum which could also be a sign of dark matter pair annihilation. The agreeing measurements of the antiproton-to-proton flux by the PAMELA telescope [64] and BESS (Polar Antarctic flight) [65] did not show any excess, but could on the other hand indicate new constraints on the mSUGRA parameter space [66]. In contrast, the positron spectrum shows a clear excess between 10 GeV and 100 GeV [67]. Similar observation was reported by the ATIC Balloon Experiment [68]. While the excess could be interpreted as an indirect dark matter observation, there are also other strong evidence that the result can be explained by other sources like pulsars [69].

Despite of the extensive searches for SUSY, both within particle- and astrophysics, no experiments have so far reached any conclusive observation. The question of SUSY remains open as we prepare for the LHC experiments to explore new and very exciting energy regimes.

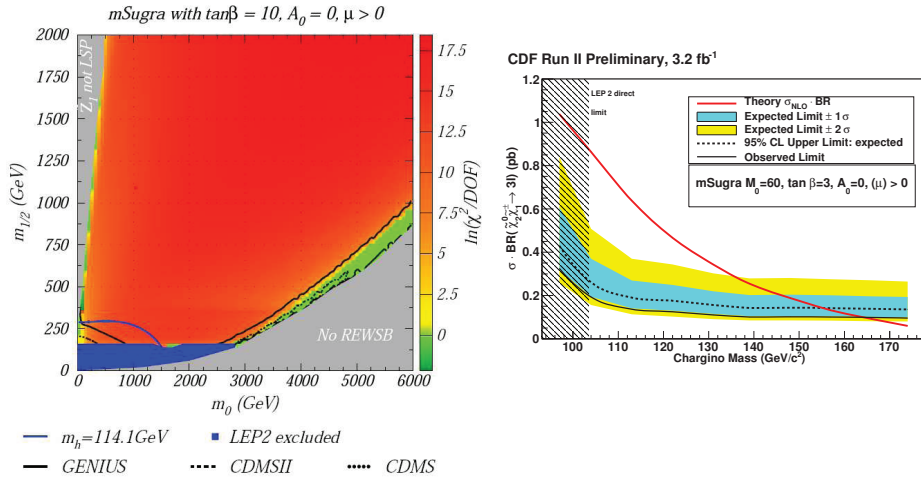


Figure 3.4: Left: Plot of  $\chi^2/\text{dof}$  for the mSUGRA model [51]. Right: expected and observed limit for the mSUGRA model ( $m_0 = 60 \text{ GeV}$ ,  $\tan\beta = 3$  and  $A_0 = 0$ ). The expected limit of about 156 GeV, and observed limit is of 164 GeV [58].

# Chapter 4

## The experimental challenge

### 4.1 The LHC accelerator

This section will describe the new generation of the experimental apparatus at CERN. The main goal of the LHC experiments is to finally pin down the mechanism of the electroweak symmetry breaking with a discovery of the Higgs particle. In addition, there is a rich scientific program for consolidation of the Standard Model as well as searches for extensions of it.

The colliding particles and their energy are the raw material for producing new particles. Whichever particles one chooses to collide, there are always both technical and experimental advantages and disadvantages. Colliding leptons which are elementary particles like at LEP ( $e^+e^-$  collider) makes it possible to control precisely the center-of-mass energy. Such conditions are favourable for precision measurements. The two beams of LEP could circulate in the same vacuum pipe and be bent by the same dipole field. But acceleration of light particles like electrons is strongly limited by the energy loss due to synchrotron radiation.

Hadrons, on the other hand, are composite particles and the hard processes of interest occur between their constituents, quarks and gluons, which carry only a fraction of the total energy. The resulting significantly varying center-of-mass energy is suitable for discoveries of unknown particles and phenomena as it allows to perform wide energy scans. At the same time, the LHC is expected to deliver a very high number of collisions which also make it possible to do several precision measurements. In addition, the relatively large mass of the proton compared to electrons reduces the synchrotron radiation energy loss and allows them to be accelerated to very high energies.

The two very important features of an accelerator are (1) the center of mass energy, the energy that can go into the processes and sets the scale of the experiment, and (2) the intensity of the beam which in turn gives the number of collisions that one can expect. LHC will collide protons and when setting the beam energy one has to take into account that only a fraction of the total energy carried by the proton is available. At the nominal energy of 7 TeV per beam, the average parton center-of-mass energy will be around 1 TeV. This choice is motivated by the theoretical estimates of the plausible energy scale of some of the theories we wish to probe.

In order to maximize the collision rate the choice has been to collide protons as today's anti-proton production techniques are not efficient enough and would be a limiting factor like at the Tevatron. At the same time, in collisions of 7 TeV protons it is the gluon

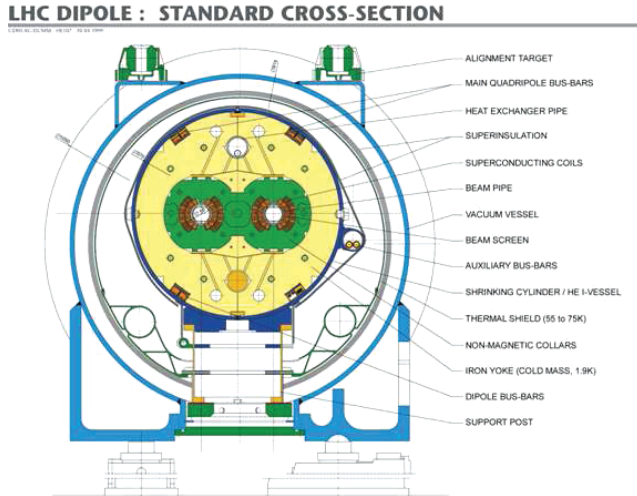


Figure 4.1: Cross section of a dipole magnet showing the two-in-one design.

interaction that will dominate the hard processes and the advantage of colliding anti-protons would be small.

Based on the theory we can calculate cross sections for interesting processes which are a measure of the probability with which these occur. The number of expected events of a particular type can then be estimated by multiplying the cross section with the integrated luminosity delivered over some period of data taking. The luminosity is defined by the beam parameters [70]:

$$L = \frac{f_{rev} n_b N^2}{\sigma_x \sigma_y} F(\Phi, \sigma_{xy}, \sigma_s) \quad (4.1)$$

where  $\sigma_x$  and  $\sigma_y$  are the transverse root mean square (r.m.s.) beam sizes at the interaction point;  $f_{rev}$  is the revolution frequency;  $n_b$  the number of particle bunches;  $N$  the number of particles in each bunch; and  $F$  is a reduction factor that depends on the crossing angle ( $\Phi$ ) between the beams, the transverse r.m.s. beam size ( $\sigma_{xy}$ ) and  $\sigma_s$ , the r.m.s. beam bunch length.

The LHC design luminosity  $L = 10^{34} \text{ cm}^{-2} \text{ sec}^{-1}$  will be achieved with 2808 bunches per beam each containing  $1.15 \times 10^{11}$  protons. The transverse r.m.s. bunch size at the interaction point is  $16 \mu\text{m}$  while the longitudinal size is  $7.5 \text{ cm}$ . The crossing angle  $\Phi$  is  $320 \mu\text{rad}$ . With a bunch time spacing of  $25 \text{ ns}$  the bunch frequency  $f_{rev}$  will be  $40 \text{ MHz}$ . Under these conditions there will be approximately 23 minimum bias interactions per beam crossing which will come on top of the interesting hard scatterings.

#### 4.1.1 The LHC magnets

The choice of two proton beams requires two separate vacuum pipes with a bending magnetic field of opposite polarities. As LHC is using the tunnel which was built for LEP, the radius of  $4.3 \text{ km}$  is given. Trying then to achieve as high energy as possible, one needs not only a sufficient accelerating facility, but also a strong enough magnetic field. The 7-TeV protons need a bending field of  $8.4 \text{ T}$ . In order to achieve this within the available space in the tunnel, LHC has developed compact superconducting magnets. These are based on

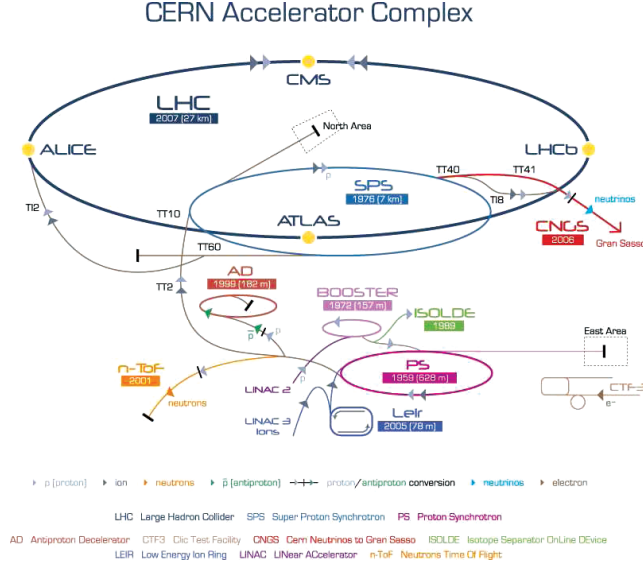


Figure 4.2: The CERN accelerator complex.

a novel two-in-one design where the two magnetic coils share the same infrastructure and cryostat, see Figure 4.1. In order to reduce the number of interconnections the magnets are as long as 15 m. The superconducting coils which are made of niobium-titan and cooled by super-fluid helium at the operation temperature of 1.9 K, carry a current density of  $1.5\text{--}2\text{ kA mm}^{-2}$ . Operating at such extreme conditions, the magnets are equipped with advanced monitoring and quenching protection systems.

The LHC is using the concept of strong focusing applying dedicated quadrupole magnets to prevent divergent trajectories, reduce the number of transverse oscillations per revolution and thus squeeze the beam. In addition there is a whole system of focusing magnets in front of the collision points that reduces the transverse size of the bunches.

The focusing has an additional effect. Together with two dedicated collimators it contributes to preventing stray particles which otherwise could harm the machine elements and the detectors.

### 4.1.2 The LHC accelerator complex

The LHC storage ring is only the last step in a long acceleration chain. Figure 4.2 shows an overview of the CERN accelerator complex. The system reuses machines constructed during the 50 years of CERN history. Protons achieved from hydrogen plasma are first accelerated in a linear accelerator (LINAC2) and then passed into the Booster synchrotron for acceleration up to 1.4 GeV. The Proton Synchrotron (PS) brings it up to 26 GeV. While in the PS, the beam is split into bunches with appropriate spacing needed for the LHC. After a 7 km transfer line the beam is accelerated in the Super Proton Synchrotron (SPS) up to 450 GeV. Then follows a technically challenging injection into the LHC ring where the beam is gradually brought up to the nominal energy of 7 TeV. In addition to protons, LHC will also have dedicated runs colliding lead ions up to 5.5 TeV.

Once the data taking starts the beam will gradually loose intensity as the number

of particles is reduced by the physics collisions and collisions with beam gas and other distortions along the trajectory. The turnaround time before a new beam is established is approximately 5 hours. So in order to achieve efficient operation, the optimal physics run time will be around 10 hours.

During the last three years two magnet incidents have caused delays of the LHC start-up. On March 27th 2007 a support structure of one of the quadrupole magnets broke during high-pressure tests at CERN. The structures were redesigned and repaired.

On the 10th of September 2008 LHC circulated for the first time a proton beam in each direction. The testing was very successful and progressing towards the first collisions. However, on the 19th of September a serious helium blow out was caused by a defect electrical connector between two magnets which created an electrical arc and a hole in the cryostat. The pressure from expanding helium opened the relief valves, however the system was not able to contain the pressure rise and as a result of the large pressure forces, the magnets were pushed apart, some places even breaking the anchors in the concrete floor [71].

The damaged parts have been replaced or repaired and similar connectors have been checked in order to avoid the problem in the future. The relief valves have increased in number and capacity and the anchoring of the magnets to the floor has been strengthened at the locations of the vacuum barriers. Due to these works the LHC start-up was postponed until November 2009 when it for the first time collided beams at the center of mass energy of 450 GeV. After a short break from the end of December 2009 until March 2010, the beam energy has been increased up to 3.5 TeV. LHC plan to continue to collect collision data at 7 TeV in a long run lasting until the end of 2011 and the goal is to collect  $1 \text{ fb}^{-1}$ . LHC plans then to shut down in 2012 in order to repair the magnet splices. This improvement is necessary in order to operate the machine safely at the design energy and luminosity.

Depending on the performance of the LHC and the physics results, there are plans for an upgrade to the so-called Super LHC. The first step will be to increase the luminosity to  $10^{35} \text{ cm}^{-2} \text{ sec}^{-1}$ .

## 4.2 ATLAS detector

The LHC machine is designed to deliver proton beams with the energy and luminosity required to produce the heavy unstable particles related to the interesting rare processes of new physics at a measurable rate. There are four detectors around the accelerator, two specialized, ALICE for heavy ion physics and LHCb for study of  $b$ -quark systems, and two general purpose detectors: ATLAS [72, 73] and CMS. Related to ATLAS there are three small detectors in the very forward region specializing in measurements of luminosity and soft proton collisions. These will be briefly described in Section 4.2.5. The TOTal Elastic and diffractive cross section Measurement (TOTEM) is a similar forward detector associated with CMS.

The design of the two general purpose detectors is driven by their main focus on the discovery of the Higgs, new physics, but also on SM measurements. The LHC experimental conditions dictate very demanding technical detector standards as they have to be very fast and radiation tolerant. At average the detectors will be filled with approximately 1000 particles per collision, every 25 ns, which requires very high granularity, specially in the innermost layers and at the same time close to full solid angle coverage in order to



record the entire event.

Since the LHC is basically a “gluon collider”, QCD processes with a total cross section of approximately 100 mb will be absolutely dominant. The interesting processes can only be spotted based on signatures that can suppress the QCD background like leptons, photons, missing transverse energy, displaced vertices or remarkably energetic jets. The detector functionality is therefore optimized for reconstruction of such benchmark discovery signatures.

ATLAS is a cylindrical symmetric detector which can roughly be divided in three parts: the central barrel and two end-caps. The choice of the magnet system layout sets the impressive scale of the apparatus: it is 25 m high and 44 m long.

The ATLAS coordinate system is a right-handed system: the  $z$ -axis points along the beam in the counterclockwise direction, the  $x$ -axis points from the interaction point towards the center of LHC, while the  $y$ -axis points upwards. The frequently used cylindrical coordinates are:  $\phi$  – azimuthal angle around the beam line,  $R$  – the radius and  $\theta$  – the polar angle with respect to the beam line. However, the pseudo-rapidity defined in the relativistic limit as

$$\eta = -\ln[\tan(\theta/2)] \quad (4.2)$$

is an even more natural choice than  $\theta$  as the light particles are produced such that  $d\sigma/d\eta$  is constant. Another advantage is that the pseudo-rapidity difference is invariant under Lorentz boosts along  $z$ .

As LHC collides composite particles the precise center-of-mass energy remains unknown. However, the initial momentum in the transverse plane is negligible. The picture of the event in the transverse plane becomes therefore important and the frequently used transverse quantities are defined as:

$$p_T = \sqrt{p_x^2 + p_y^2}, \quad E_T = \sqrt{E_x^2 + E_y^2} \quad (4.3)$$

Summing up the  $E_T$  of an event one can calculate the transverse energy deficit  $\cancel{E}_T$ . The angular spacing is defined as:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \quad (4.4)$$

Figure 4.3 shows how ATLAS is built up from many subdetectors. The innermost layer (Figure 4.3: yellow central part) measures the tracks of charged particles down to a momentum of  $\sim 0.5$  GeV and is made of very radiation hard and light components in order to be as transparent as possible. The granularity is adjusted to the high track density towards the interaction point.

The tracking volume is inside a relatively strong magnetic field and the precision track reconstruction provides charge identification and momentum measurement, as well as reconstruction of displaced vertices used in identification of B-mesons (D-mesons to a smaller degree) and tau leptons.

Outside the thin solenoid magnet coil enclosing the Inner Detector (Figure 4.3: thin red layer) we find the calorimeters which measure the energy of the incoming particles by stopping them. The electrons and photons which easily interact with matter and create electromagnetic showers are measured in the innermost electromagnetic calorimeter (Figure 4.3: green). The heavier hadrons which easier penetrate matter and create more longitudinally stretched hadronic showers are stopped in the thicker and more dense hadronic calorimeter (Figure 4.3: outer red layer). The only particles that escape the

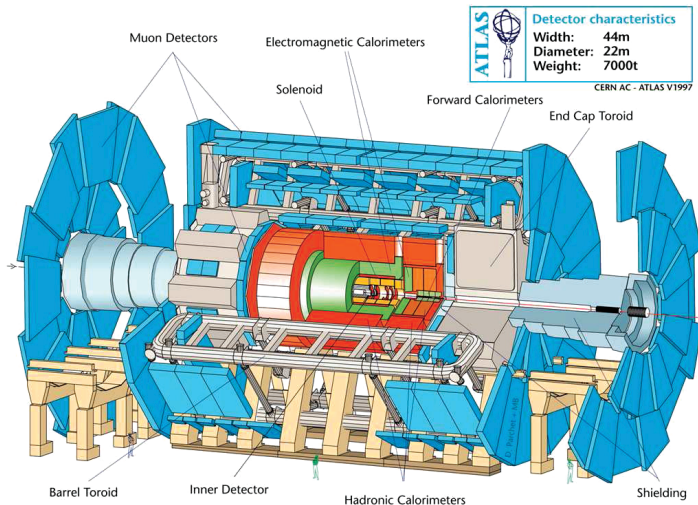


Figure 4.3: The ATLAS detector with one quarter removed in order to show the cross section through the various subsystems.

calorimeters, with the obvious exception of the weakly interacting particles, are the muons. The outermost layers (Figure 4.3: blue plates) are dedicated to precise measurement of their tracks, and thus also momentum and charge. In addition they provide the very important fast trigger information as muons are a typical signature of many interesting physics scenarios.

Except for the original toroid magnet layout (Figure 4.3: gray) and impressive size, the ATLAS structure is rather conventional. However, the actual technologies involved are at the cutting edge of the development trying to face the extreme conditions of the LHC collision environment. The following sections will give a more detailed description of each of the subsystems.

### 4.2.1 The Inner Detector

The Inner Detector (ID) consists of three independent but complementary tracking sub-detectors. The layers close to the interaction point are designed to tolerate the highest radiation doses in the detector. The large track density also requires very fine granularity and resolution as well as fast response. Figure 4.4 shows the different subdetectors and their size and pseudo-rapidity coverage.

The two innermost layers are the Pixel and Silicon microstrip detector (SCT) trackers which provides precision track measurements. The third and outermost layer is the Transition Radiation Tracker (TRT).

In order to measure the momentum of the charged particles the Inner Detector is placed inside a superconducting solenoid magnet which is 5.3 m long, 2.5 m in diameter and provides a 2-T field along the  $z$ -axis. The ID system provides tracking information up to  $|\eta| < 2.5$  and the momentum resolution is [73]:

$$\sigma_{p_T}/p_T = 0.05\%p_T \oplus 1\%. \quad (4.5)$$

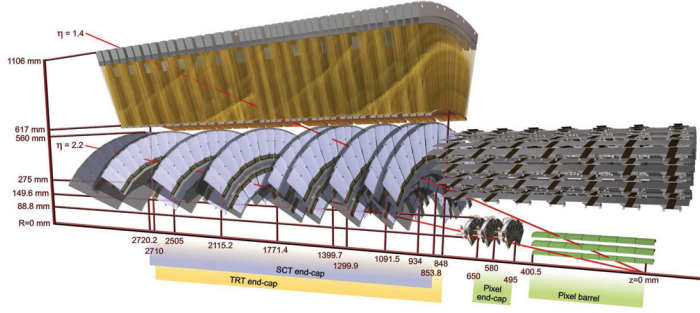


Figure 4.4: The ATLAS Inner Detector. [73]

### Pixel Detector

The Pixel detector consists of three layers of pixel sensors, each with the external dimensions of  $19 \times 63 \text{ mm}^2$  mounted on a very light cylindrical support structure and arranged in three barrel layers and two end-caps with three discs in each.

The minimum pixel size is  $50 \times 400 \text{ }\mu\text{m}^2$  and as there are 47232 of them on each sensor, this represents a major challenge both in terms of electronics, readout and cooling (see Section 4.4). The pixel detector has approximately 80.4 million readout channels and the intrinsic accuracies are  $10 \text{ }\mu\text{m}$  in  $(R - \phi)$  and  $115 \text{ }\mu\text{m}$  in  $R$  direction, both in the barrel and the end-caps.

### The SemiConductor Tracker - SCT

SCT is the middle-layer subdetector of the ID and is designed to provide precise measurements over a large volume. As the track density is decreasing with larger radius, SCT uses silicon strip sensors with coarser granularity than in the pixel detector. This choice reduces the costs and number of readout channels. Each SCT module is double sided and consists of four  $63.6 \times 64.0 \text{ mm}^2$  silicon wafers with  $80 \text{ }\mu\text{m}$  pitch micro-strips and a stereo angle of  $40 \text{ mrad}$  to improve the  $z$ -measurement. The barrel SCT has four cylindrical layers while the two end-caps have nine discs.

The intrinsic accuracies are  $17 \text{ }\mu\text{m}$  in  $(R - \phi)$  and  $580 \text{ }\mu\text{m}$  in  $R$  direction. Although the SCT covers a much larger volume than the pixel detector the number of readout channels is down to 6.3 millions, which is still a considerable number.

### Transition Radiation Tracker - TRT

The TRT is made out of straw tubes with a diameter of  $4 \text{ mm}$ . These are tightly packed in layers which are interleaved with foils; 73 layers in the barrel and 160 straw planes in the end-caps.

The straws are filled with a Xenon-based gas and a thin anode wire stretched through the center of the straw collects the ionization charges. The interleaving material provides transition radiation which starts at  $\gamma \sim 1000$  where  $\gamma = E/m$ . The light electron will emit transition radiation already at  $\sim 1 \text{ GeV}$  while heavier particles like the  $\pi$ -meson first at  $\sim 100 \text{ GeV}$ . Due to these different thresholds TRT can provide  $e/\pi$  separation up to approximately  $100 \text{ GeV}$ .

The TRT provides measurements within  $|\eta| < 2.0$  and only in the  $R - \phi$  plane. A typical track has approximately 36 hits which, once reconstructed, give an almost bobble chamber like track picture with a track accuracy of  $130 \mu\text{m}$ .

## 4.2.2 Calorimetry

The calorimeter system is placed outside the ID and the solenoid magnet. One of the important constraints during the construction of the ID was to introduce as little material as possible in front of the calorimeter. Especially electrons which easily emit Bremsstrahlung and photons which undergo conversion are seriously affected. These disturbing processes spoil in many ways both the track reconstruction and the energy measurements. In the final setup the total mass in front of the calorimeters adds up to  $\sim 4.5$  tons, which in some regions represents as much as  $\sim 2 X_0$  or  $\sim 0.6\lambda$ <sup>1</sup>.

The different nature of the electromagnetic and hadronic interactions in matter require the more fine-grained electromagnetic calorimeter to be the innermost calorimeter layer. It is then surrounded by a more dense hadronic calorimeter with a coarser granularity. The granularity and structure are optimized to support both particle identification and trigger information. Energy measurements impose additional important design requirements: full coverage in eta, with forward components covering up to  $|\eta| < 4.9$ , and sufficient longitudinal depth in order to prevent punch-through into the muon system. The hermeticity requirement is especially important for a reliable measurement of the energy deficit ( $\cancel{E}_T$ ) which is one of the important signatures of many new processes. Table 4.1 gives an overview of the calorimeter system performance goals and  $\eta$ -coverage.

| Detector component          | Required resolution                       | $\eta$ -coverage     |                      |
|-----------------------------|---|----------------------|----------------------|
|                             |   | Measurement          | Trigger              |
| EM Calorimeter              | $\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$ | $\pm 3.2$            | $\pm 2.5$            |
| Hadronic calorimetry (jets) |   |                      |                      |
| barrel and end-caps         | $\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$   | $\pm 3.2$            | $\pm 3.2$            |
| forward                     | $\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$ | $3.1 <  \eta  < 4.9$ | $3.1 <  \eta  < 4.9$ |

Table 4.1: ATLAS calorimeter system performance goals [73]. The unit for  $E$  is GeV.

Figure 4.5 shows a cross section of the ATLAS calorimeter system. The various parts will now be described in more detail.

### LAr electromagnetic calorimeter - ECAL

The barrel EM calorimeter is 3.2 m long, has an inner diameter of 2.8 m, an outer diameter of 4.0 m and covers  $|\eta| < 1.475$ . It is constructed as two identical half-barrels separated by a small 4 mm gap at  $z = 0$ . The end-cap calorimeters consist of two coaxial wheels covering  $1.375 < |\eta| < 2.5$  and  $2.5 < |\eta| < 3.2$ . The performance of the ECAL in the transition region between the barrel and the end-caps,  $1.37 < |\eta| < 1.52$ , is degraded due to the large amount of material in front and therefore often excluded from analysis. The depth of the ECAL is estimated to be at least  $\sim 22 X_0$ .

ATLAS has chosen liquid Argon (LAr) as the active material due to its radiation hardness and good performance in terms of energy resolution. The absorbers consist

<sup>1</sup> $X_0$  stands for the electromagnetic radiation length while  $\lambda$  is the nuclear interaction length.

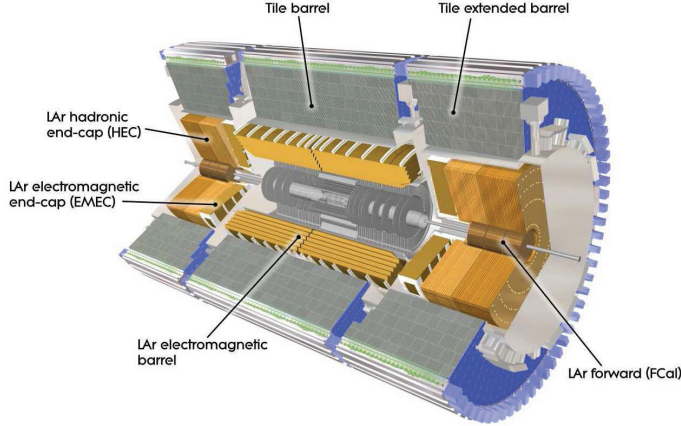


Figure 4.5: The ATLAS calorimeter system includes a LAr electromagnetic calorimeter and a hadronic calorimeter. [73]

of accordion shaped lead plates with kapton electrodes, a geometry that allows full  $\phi$  symmetry without azimuthal cracks for readout.

In order to reduce the amount of material in front, ECAL shares the vacuum vessel with the solenoid and is operated at 89 K. A LAr pre-sampler covering  $|\eta| < 1.8$  is placed in front of the ECAL in order to measure the  $e/\gamma$  energy losses occurring before the calorimeter is reached. The innermost layers of the ECAL are also designed with a finer granularity in order to improve  $\pi^0/\gamma$  separation,  $\gamma$ -conversion reconstruction and isolation energy measurements [74].

### Hadronic calorimeter - HCAL

The outer calorimeter layer is dedicated to measurements of the energy of hadronic particles over as large  $\eta$ -range as possible. The barrel HCAL is 5.8 m long and has an inner and outer radius of 2.28 m and 4.25 m respectively. It covers an  $|\eta|$  range up to 1.7 where the 2.6 m long extended barrel modules take over. The barrel HCAL is a sampling calorimeter using steel as absorber and scintillating tiles as the active material, arranged in a vertical tile geometry which facilitates for an easier light transfer. The depth of the HCAL at  $\eta = 0$  is  $\sim 10 \lambda$ .

**LAr hadronic end-cap calorimeter - HEC** is placed outside the ECAL end-cap wheels and covers the range  $1.5 < |\eta| < 3.2$ . As the radiation doses increase in the more forward regions, the materials chosen for this component are the radiation hard LAr and copper plates. The outer radius is 2.03 m while the inner radius reaches as close as 0.475 m to the beam pipe.

**LAr forward calorimeter- FCal** is the innermost calorimeter layer designed to sustain the intense forward particle flux in the region  $3.1 < |\eta| < 4.9$ . The harsh environment in this location combined with very limited longitudinal space calls for a very dense and radiation hard material. Again LAr is chosen as the active material. The first module is made of copper optimized for electromagnetic showers, while the last two measuring

mainly the energy of hadronic interactions use tungsten as absorber .

### 4.2.3 The magnet system

The ATLAS magnets are a hybrid system consisting of four large modules: the inner solenoid, the barrel toroid and two end-cap toroids. The size of the toroid system enhances the precision of the muon momentum measurement by providing a large lever arm while the air core reduces multiple scattering.

The Inner Detector is enclosed inside a superconducting solenoid magnet which is 5.3 m long and 2.5 m in diameter. It is operated at 4.5 K and a current of 7.730 kA provides a 2 T field along the  $z$ -axis. One important design requirement has been to make the coil as thin as possible in order not to spoil the energy resolution. The final solenoid assembly represents  $\sim 0.66$  radiation lengths. The steel in the Hadronic Calorimeter returns the flux.

One of the most spectacular features of the ATLAS detector is the outer toroid magnet system. The light air-core structure reduces multiple scattering facilitating for a large tracking volume with relatively strong magnetic field. There are three toroid magnets. The large barrel magnet with eight  $\sim 25 \text{ m} \times 5 \text{ m}$  coils, each in its own cryostat, are held together with an aluminium mechanical structure. The two end-cap toroids, also with eight coils in each, are contained in a common cryostat. The coils are made of aluminum stabilized NbTi super conductor. They operate at 4.6 K with a current of 21 kA generating a local field of 4 T in the windings. The field is strongly varying but results in an average strength of 0.5 T in the barrel (corresponding to  $|\eta| < 1.4$ ) and 1 T in the end-caps ( $1.6 < |\eta| < 2.7$ ). The transition region  $1.4 < |\eta| < 1.6$  between the barrel and the end-caps is covered by a combination of fields from the two modules.

### 4.2.4 The Muon Spectrometer

From the interaction point to the outer radius of the calorimeters, the material of the detector adds up to  $\sim 11 - 15 \lambda$ . The only particles surviving that far, with the exception of the weakly interacting ones, are the muons which at typical LHC energies are minimum ionizing particles. Their track is registered in the tracking system (ID), they leave small energy deposits in the calorimeters, but their main signature are hits in the muon spectrometer. A muon needs a minimum  $p_T \sim 3 \text{ GeV}$  in order to reach the spectrometer. Figure 4.6 shows a cross section of the muon spectrometer installed together with the toroid magnet.

In the barrel region the muon spectrometer consists of three planes of chambers at radii of 5, 7.5, and 10 m. This spectrometer region covering up to  $|\eta| = 1.4$  is immersed in the barrel toroid magnetic field which provides a bending power  $\int B dl$  of 1.5-5.5 Tm. In the more forward region there are four disc shaped layers of chambers at  $|z|$  equal 7.4 m, 10.8 m, 14 m and 21.5 m, covering the pseudo-rapidity region of  $1 < |\eta| < 2.7$ . Here the main source of the magnetic field is the end-cap toroid which provides 1-7.5 Tm of bending power in  $1.6 < |\eta| < 2.7$ . The overall spectrometer layout is such that a muon will typically traverse 3 muon chamber layers as shown in Figure 4.6.

The muon spectrometer faces two tasks: providing a precise track and momentum measurement and a fast trigger based on muon momenta down to 5 – 10 GeV. The system is therefore assembled from two complementary types of chambers.



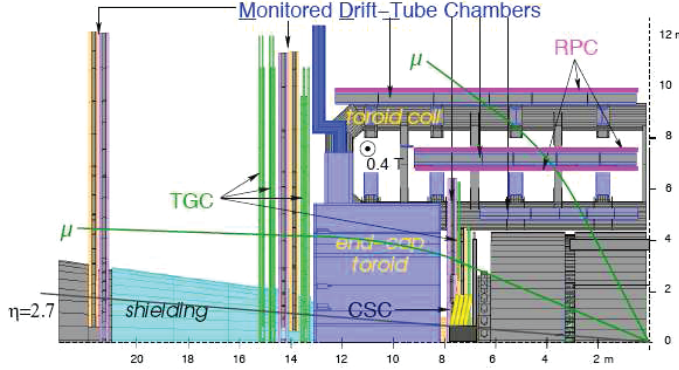


Figure 4.6: A cross section of the The ATLAS muons spectrometer where the four different chamber technologies are indicated by color. MDTs and CSCs provide precision tracking while the fast RPCs and TGCs also provide trigger information in addition to position measurements. [73]

For the precision measurement ATLAS uses Monitored Drift Tubes (MDTs). The tubes are made of aluminium with a diameter of 30 mm and a length varying from 85 cm to 650 cm. They are filled with an argon-based gas at an absolute pressure of 3 bar and equipped with a central anode wire which collects the electrons created by the passing muons. The tubes are arranged in modules with 3-8 layers in each and this type of chambers is used in the whole  $\eta$  range.

The innermost forward precision measurement are provided by the Cathode-strip chambers (CSCs) placed at  $|z| = 7.4$  m, covering  $2 < |\eta| < 2.7$ . These are multi-wire proportional chambers with cathode strips perpendicular to the wires as well as parallel to the wires providing the transverse coordinate. The forward disc modules have four CSC planes.

The precision chambers are relatively slow compared to the 25 ns beam crossing time during which a particle with typical momentum travels not more than 7.5 meters, a distance which is shorter than the radius on the ATLAS detector. The trigger chambers are therefore optimized for excellent time resolution and fast response. Furthermore, they are required to cover  $|\eta| < 2.4$  and the whole azimuthal angle  $\phi$ . Given the wide  $\eta$  range, the  $p_T$  resolution requirements are difficult to meet in the most forward region as the momentum  $p$  is there very large for a given  $p_T$ . The granularity of the forward modules has therefore to be sufficiently increased in order to match the barrel performance. In addition, the end-caps must be capable of handling a 10 times larger radiation dose than the barrel. Consequently, two different technologies have been applied.

As trigger layers in the barrel, ATLAS uses Resistive Plate Chambers (RPC) which consist of gaseous parallel electrode plates that collect the avalanche charges. RPCs provide both good spatial and time resolution as well as a sufficient high-rate capability. The end-caps trigger layers use the so called Thin Gap Chambers (TGCs), a kind of multi wire proportional chambers devices. This technology has been selected due to the good time resolution and hit rate capability. Both types of trigger chambers can deliver a signal within 15-25 ns from the collision and are capable of providing a beam crossing tag.

The required resolution for the muon system is

$$\sigma_{p_T}/p_T = 10\% \quad \text{at} \quad p_T = 1 \text{ TeV}. \quad (4.6)$$

The overall performance of the muon system requires in addition very precise alignment information which is provided by an optical system in combination with a track based alignment. In order to achieve the planned resolution one needs to know how the chambers are positioned with a precision of 30  $\mu\text{m}$  and perform detailed monitoring of the magnetic field.

### 4.2.5 ATLAS forward detectors

ATLAS is equipped with three detectors in the very forward region mainly dedicated to measure the delivered luminosity. LUCID (LUMinosity measurement using Cerenkov Integrating Detector) detects inelastic  $p$ - $p$  scattering. It is placed at  $\pm 17$  m from the collision point and is the main online luminosity monitor. Further down the beam line is the ALFA detector (Absolute Luminosity For ATLAS). It is based on scintillating fiber trackers inside Roman pots which allow it to come as close as 1 mm to the beam.

The third is the Zero-Degree Calorimeter (at  $\pm 140$  m from interaction point) and it is used to determine the centrality of heavy-ion collisions.

## 4.3 ATLAS Trigger and Data Acquisition - TDAQ

The power of LHC lies not only in the tremendous center-of-mass energy of the colliding protons, but also in the high design luminosity. However, this represents also a great challenge in terms of data acquisition. As it is technically entirely impossible to record all events, the LHC experiments must develop sophisticated trigger systems which would reduce the rate of events to be stored to a feasible amount, while at the same time selecting the interesting ones. For this purpose ATLAS has implemented a three-level trigger system.

The starting point is a bunch crossing rate of 40 MHz and an interaction rate of  $\sim 1$  GHz. The Level 1 trigger (L1) receives the data from the pipeline memory and looks for high- $p_T$  muons, electrons, photons, jets, decay products of taus and also large missing and total transverse energy. It is a hardware based trigger which uses only a fraction of the information from the subdetectors and a very coarse granularity. It has to make a very fast decision, within 2.5  $\mu\text{s}$ , and at the same time create so-called Regions-of-Interest (RoIs). RoIs are geographical coordinates in  $\eta$  and  $\phi$  of the detector regions where the interesting objects were found. The implemented selection criteria are called “trigger menus” and represent some characteristic signatures based on particles and objects present in an event passing certain thresholds. Prescaled menus which only let through a fraction of events satisfying a certain trigger requirement are also available.

If the L1 trigger accepts an event, it is processed by the detector-specific front-end systems, the Read Out Drivers (RODs), and passed to the Level 2 trigger (L2) at the L1 reduced rate of  $\sim 75$  kHz. Now the L2 trigger uses the full granularity and precision of the detector, but only in the RoIs which represent approximately 2% of the data. L2 is a software-based trigger and it reduces the event rate to  $\sim 3.5$  kHz with a latency of  $\sim 40$   $\mu\text{s}$ .



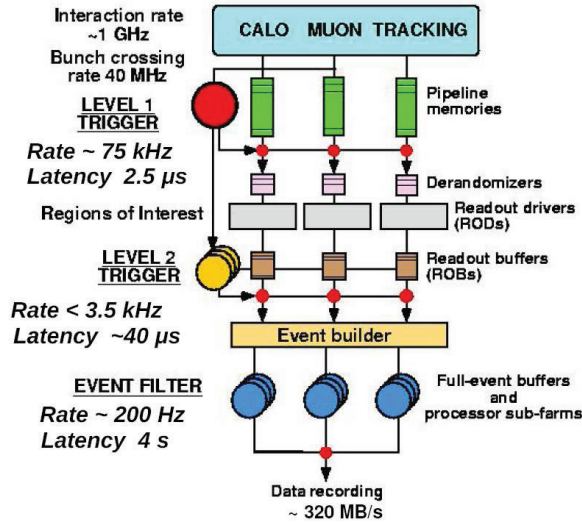


Figure 4.7: The ATLAS Trigger and Data Acquisition System. From [73] with modifications.

If an event is accepted at L2, it is passed to the event building system and further to the last trigger level: the Event Filter (EF). The EF is also a software based trigger and together with the L2 it is considered as the High Level Trigger (HLT). The EF uses the full detector information and the offline analysis procedures including track fitting, vertexing and missing  $E_T$  calculations. The processing time is around 4 s per event and it is performed by the Point 1<sup>2</sup> computing farm. The EF reduces the event rate to  $\sim 200$  Hz which with an event size of  $\sim 1.6$  MB represents a data writing rate of 320 MB/s. Events accepted by the EF are transferred from the Point 1 facility to the central CERN computing center, the so called Tier 0. Figure 4.7 shows a schematic overview of the parallel triggering and data acquisition process.

In addition to the acquisition of data, TDAQ collects and stores the configuration, control and monitoring information of the hardware – the detector – and the software components. Based on this information one will be able to set data quality flags and understand better the reliability of the data.

## 4.4 ID cooling

Despite the high radiation doses, the silicon detectors SCT and the Pixel, with exception of the innermost layer, are foreseen to last through the whole LHC operation time. The most severe problems are related to silicon bulk radiation damages which cause changes in the effective doping, leakage currents and noise. One of the consequences is a change in depletion and operation voltage. In order to reduce the leakage currents, silicon detectors are operated at a temperature of  $-7^\circ\text{C}$  provided by a common cooling system. At the design luminosity it has to remove  $\sim 60$  kW of heat [75].

For this purpose ATLAS has installed an evaporative cooling system. Compared to a

<sup>2</sup>Point 1 is the name of the ATLAS experiment site at CERN.

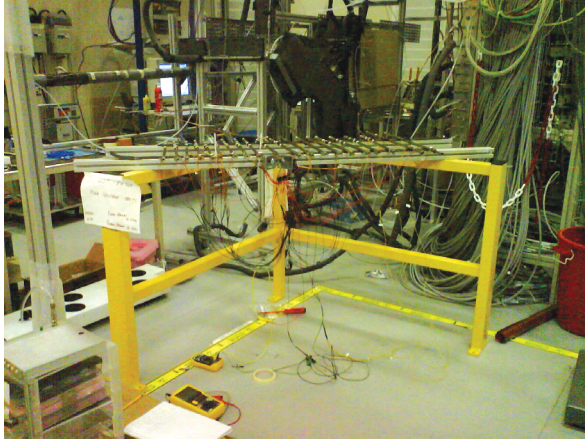


Figure 4.8: Setup for heater pulse tests.

mono-phase system it has the advantage of high heat transfer rate, smaller temperature gradients along the cooling channels and small size of the channels. The required cooling is thus obtained introducing a minimum of additional material.

The working principle is similar to what is used in refrigerating industry. The  $C_3F_8$  coolant is distributed by four racks in the cavern into an intricate system of 4 mm (inner diameter) copper pipes. The distribution system consists of 204 individual circuits. The coolant enters at room temperature in liquid form. As it passes small-diameter capillaries, the pressure drops. The coolant exits the capillaries in a two-phase form and starts to boil lowering the temperature of the cooling structure attached to the detector modules. Exiting the ID system, the two phase fluid must be heated up to the dew-point of the cavern in order to avoid condensation on the exhaust pipes. This is done by special heaters powered with 110 V pulses in order to maintain a constant exhaust gas temperature of 20°C.

The strongly varying temperatures of the ID require it to be flushed with dry nitrogen in order to avoid condensation. Due to the difficult conditions, the ID environment is subject to continuous detailed monitoring and secure operation conditions are maintained by the Detector Control System (DCS).

During the ATLAS commissioning the ID cooling system suffered from two major incidents. In February 2007 a serious problem surfaced as one of the heaters exploded due to a short circuit to ground. It was realized that a number of heaters had a too low isolation resistance. The problem was related to the materials used in the connectors and the fact that there are small amounts of water in them. Therefore, the heaters were equipped with sleeves which keep the moisture out. This did not solve entirely the problem and the heaters had to be redesigned. 120 heaters had to be moved to locations which were accessible at a later stage in the commissioning.

Before the heaters could be reinstalled their isolation resistance was tested by longer periods (typically several days) of electrical pulses in order to see if the resistance stayed sufficiently high and stable. Figure 4.8 shows the pulse test setup.

The re-engineering of the pipe work was finished towards the end of 2007 and heaters with redesigned connectors were finally in place during the first months of 2008.

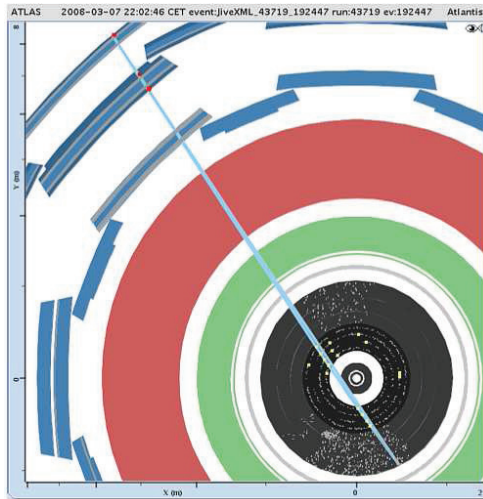


Figure 4.9: Combined track during the Milestone 6 (M6) cosmic run.

The SCT cooling was back on track in February 2008 for an SCT-technical run. In March followed one of the so-called ATLAS Milestone cosmic runs, sixth in the row (M6), with the cooling in stable 24 hour operation. During M6, SCT ran for the first time in a combined ATLAS run with the TRT, scintillator triggers and the muon RPCs. Figure 4.9 is from M6 and shows a cosmic muon passing the RPCs, TRT and traversing the whole barrel SCT leaving 8 space points<sup>3</sup>.

A second incident struck the cooling system in May 2008 and caused considerable delays. Three out of six compressors were damaged due to a burn out which caused a leakage of 100 kg of the coolant liquid, while the remaining was contaminated. The cooling plant was repaired and the coolant and dirty pipe-work replaced. Additional filters and sensors were installed in order to avoid similar problems in the future. The cooling was successfully reestablished for the beam-pipe bake-out (summer 2008). The bake-out process established a high vacuum in the beam pipe by heating it up to 220 C from the outside. A well functioning cooling system was crucial during this procedure, especially for the innermost  $b$ -layer of the pixel detector.

## 4.5 ATLAS commissioning and readiness

Secondary cosmic radiation which at sea level mainly consists of muons is an excellent natural particle source well suited for testing the detector. The flux is 130 Hz/m<sup>2</sup> for muons with  $E_\mu > 1$  GeV. The average energy is  $\sim 4$  GeV. In the ATLAS detector this corresponds to a flux of  $\sim 4$  kHz in the full muon fiducial volume.

Commissioning using cosmic rays started already in 2005 and was carried out in parallel with the ATLAS detector assembly. As the subsystems were installed they were tested both in standalone mode and in combined runs. During spring and summer 2008 there

<sup>3</sup>During a stay at CERN Autumn 2007 and Spring 2008 the author of the thesis contributed to the pulse testing of the heaters, while during M6 she did cooling and Data Quality shifts.

was a hectic testing activity preparing for the start-up of the LHC in September 2008. The main purpose of the cosmic ray commissioning has been to “time in” the L1 hardware trigger, test procedures for alignment and calibration, find dead zones or noisy channels and by this understanding to improve the detector performance. Last but not least, the ATLAS physicists could gain experience in running the Data Acquisition system (DAQ), data quality monitoring as well as the full analysis chain.

During the few days of LHC operation ATLAS recorded both so called beam splash events, where the beam was on purpose targeted at the closed collimators 140 m away from the detector, as well as beam-halo events. The beam splash events in which the whole detector was illuminated by a shower of particles were useful for checking and correcting the time calibration of the full detector. Both beam splash and beam halo events provide almost horizontal muons that were exploited to check the timing in the TRT and in the calorimeters.

After the LHC incident on the 19th of September the commissioning and calibration with cosmic rays continued during Fall 2008 with a long combined data-taking collecting 7.6 millions of tracks.

The analysis of this data showed excellent performance of the inner detector. The Pixel and SCT function with more than 99% efficiency and the track residuals comparing MC and aligned real tracks were close to perfect geometry. In general, the ID performance proved to be close to what was expected from simulations and it was considered to be in good shape for collision data.

### 4.5.1 Outlook

The LHC start-up plan was revised after the 19th of September 2008 incident and the initial plan was to ramp up to 10 TeV center of mass energy [76]. ATLAS started the preparations for a lower energy scenario by launching a large MC simulation. The analysis of these data will be shown in Chapter 9.

In November 2009 the LHC started up again and collided beams for the first time on the 23rd of November. In December 2009, after further testing, LHC delivered collisions at 900 GeV, which corresponds to the injection energy from SPS. The beams were then accelerated up to 1.18 TeV allowing to record collisions at 2.36 TeV.

During these first runs ATLAS recorded an integrated luminosity of  $20 \mu b^{-1}$  with  $12 \mu b^{-1}$  during stable beams. The results of a first look at these and a comparison of data and MC are shown in Chapter 11.

Based on the experience collected during 2009, the LHC accelerator project and the experiments revised again the plans. Considering the dangers related to accelerating the beams up to 5 TeV, it was decided that the LHC will run at 3.5 TeV for approximately two years and collect  $1 fb^{-1}$  [77]. Next there will be a shut down of approximately a year in order to improve the copper splices between the magnets before LHC can ramp up to the design center of mass energy of 14 TeV.

# **Part II**

## **Computing and Data Challenge**



# Chapter 5

## ATLAS computing

The computing requirements of ATLAS are defined by the physics searches that are the main goal of the experiment. One of the design benchmark processes is the production of the Higgs boson with a subsequent decay to leptons:  $H \rightarrow ZZ \rightarrow 4l$ . For Higgs masses between 120 and 600 GeV a cross section times leptonic branching ratio of 2.3–20.5 fb is expected [78]. At design luminosity of  $10^{34} \text{cm}^{-2} \text{s}^{-1}$ , providing the experiments approximately  $100 \text{fb}^{-1}$  of integrated luminosity per year, such events are expected to be produced, very roughly estimated, at a rate of one event per hour. This is many orders of magnitude lower than some of the dominant background processes. Important discovery scenarios like the Higgs are used to set the parameters of the experiment and the planned rate of events which will be written to storage after a careful three-level triggering filter. The last trigger level, the Event Filter (EF), will provide events at the strongly reduced rate of 200 Hz. With an expected event size of about 1.6 MB this will result in ca. 320 MB/s written to storage. The ATLAS experiment is then expected to produce an order 3.2 PB of data per year. In addition, there is also a need for Monte Carlo (MC) simulation of all processes under study which is expected to sum up to 2–3 times the volume of real data.

Both data and MC require a considerable amount of processing power and need to be distributed to a global physicist community. For their analysis, physicists will require computing power far beyond what can be provided by a desktop machine or local systems.

Facing these enormous demands and the global character of the new experiments, CERN has embraced the Grid as the model for the LHC computing infrastructure. In our daily life we are used to connect electrical devices to the power grid and utilize the energy without worrying where it comes from. The computing Grid is based on a similar idea. Users obtain a certificate and can create a proxy which gives them access to vast resources via simple universal tools. Users and institutions make their resources available on the Grid, while in return they have access to much more. The participating units can be everything from small batch farms to large clusters and advanced mass storage facilities. The computing Grid solves both the needs of processing power as well as data storage, management and movement.

This chapter will describe in more detail the Grid technology (Section 5.1) and the ATLAS data and computing model (Section 5.2). It will also describe several projects related to NorduGrid (Section 5.3) which show how the ATLAS computing challenges are solved in the area of software distribution (Section 5.6), managed production (Section 5.7) and distributed user analysis (Section 5.9). This part will necessarily contain some technical details, however, in order to make the text easier to read, they will be enclosed in

boxes which can be skipped, while more detailed information is placed in appendices.

## 5.1 Grid computing

Although lately the Grid has lost some popularity to the Cloud, it still represents some of fundamental technology behind different types of distributed computing infrastructures. Already in 1969 Len Kleinrock [79] promoted the idea of spread computer utilities with an accessibility much like the electric grid. Current Grid solutions are based on the concept defined by Ian Foster and Carl Kesselman [80]

“A computational grid is a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities.”

In order to create a computing Grid one needs a so called middleware to connect the different components like clusters and storage, manage them and provide a uniform user interface to the resources. The role of the middleware could be compared to a combination of operating system and batch system on a higher organizational level.

In a later definition Foster and Kesselam emphasized more the aspect of “coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations (VO)”. A VO is an “abstract entity grouping Users, Institutions and Resources (if any) in the same administrative domain” [81]. It can set up a set of rules for the sharing of computing resources like storage and CPUs, software and data required in a large collaborative project. Although the Grid may give users access to vast computational power, it is not suggested as an alternative to high-performance computers. Some tasks will still require low latency and high communication bandwidths which are only possible in closely coupled computers. The Grid represents rather an alternative computing model, typically focusing on harnessing of free cycles wherever available, but still in a strictly managed and controlled fashion.

In order to grasp the idea of the Grid in a more precise way Foster suggests a checklist with three points [82]:

1. Coordinate resources that are not subject to centralized control.  
The Grid is meant to include a variety of resources, from high-performance computing centers to user desktops and very different users and applications.
2. Standard, open, general-purpose protocols and interfaces should be used.  
The Grid should be built from multi-purpose protocols and interfaces in order to provide a platform which covers fundamental functionalities like user and resource management, data storage and movement.
3. The Grid is required to deliver nontrivial qualities of service.  
The resources building up a Grid should be managed and utilized in a coordinated fashion providing differentiated qualities of service regarding response time, throughput, availability, and security, and/or co-allocation of multiple resource types to meet complex user demands. The goal is a system that through combined and optimized utilization delivers more than the sum of the parts by themselves.

In the following sections which describe the ATLAS computing model and particularly in the sections concerning NorduGrid, there will be many examples of how these ideas are being implemented.



## 5.2 ATLAS Grid computing

ATLAS physicists access the resources and data through a valid Grid proxy which defines their identity and membership in the ATLAS Virtual Organization (VO). The membership information is registered in a Virtual Organization Management Service (VOMS) [83] database. When creating a proxy, users contact a VOMS server using a client tool (`voms-proxy-init`) and are granted access according to their registered role. An efficient utilization of the resources within the collaboration requires an internal hierarchy and system of privileges. Users can therefore be given three different roles, each associated with appropriate access rights and priorities. The roles are [84]:

- Grid software administrator - installs software and manages the resources
- Production manager - has a higher priority, on some systems access to more resources or queues and access to write files to common areas
- Normal users

As of 2009, ATLAS uses three Grid “flavors”: Open Science Grid (OSG) in USA [85], LCG/EGEE [86] in most of Europe, Canada and Far East and NorduGrid [87] in Scandinavia, Slovenia and Switzerland <sup>1</sup>.

Although the three Grids use different middlewares, ATLAS requires a high degree of interoperation. Therefore, on top of the basic functionalities there is a set of common tools and distributed services providing more uniform user and application interfaces.

The Worldwide LHC Computing Grid (WLCG) environment is hierarchically structured in Tiers. The CERN Computing Center is the central Tier-0 which communicates with, at the moment, 11 Tier-1 centers. Each of these Tier-1s is again a key center in a so-called “computing cloud” <sup>2</sup> enabled by one of the three Grid middlewares. The US Grid (OSG) and NorduGrid correspond each to one cloud, while the remaining 9 are all LCG/EGEE Tier-1s. A Tier-1 is considered as a high level resource dedicated for managed collaboration tasks with only restricted access for physicist users. Each Tier-1 provides therefore a number of Tier-2s and even more local Tier-3s, which are Grid enabled resources open for less organized user exploitation and analysis. The NorduGrid-related Tier-1 facilitated by the Nordic DataGrid Facility [88] (NDGF) which will be described in more detail in Section 5.3, is in its nature a distributed Tier-1 with a more flat structure and does not follow this general layout [89].

### 5.2.1 ATLAS data model

This and the following sections will describe the ATLAS data model and give a brief introduction of the different components of the computing system.

In order to better understand the ATLAS computing model, it is practical to start with introducing the different types of data files.

- RAW Data: is the output from the Event Filter (EF), the highest level of the trigger system. At design luminosity events are expected to arrive at a rate of 200 Hz and

<sup>1</sup>This list of countries involved in NorduGrid contains only those providing resources for ATLAS.

<sup>2</sup>The term “cloud” has in the WLCG context a different meaning than the more commercial concept of “cloud computing”.

mean size of 1.6 MB per event. This data in the “byte-stream” format is the input for reconstruction. It is shipped out from Point-1 to the Tier-0 in files of at most 2 GB containing events from a single run <sup>3</sup>.

- ESD (Event Summary Data): is the resulting event data containing physics objects from the reconstruction which are stored in POOL ROOT format. The content is detailed and technical enough to make a rerun of the reconstruction, so-called reprocessing, possible and thus to make access to RAW data unnecessary. The target size is  $\sim 500$  kB per event (750-900 kB per event in 2008).
- AOD (Analysis Object Data): AODs are derived from ESDs and provide an event representation suitable for analysis. The AOD physics objects should provide sufficiently detailed information for most analyses while at the same time make the event size small enough to allow the files to be widely replicated in an affordable way. The target size is set to 100 KB per event (250-290 kB per event in 2008).
- TAG (Tag Data): is the event level metadata which provides thumbnail information about the event, for example number of physics objects like electrons, jets or  $E_T$ . TAG data which is only 1 kB per event is stored in ROOT files and in relational database files to allow quick queries for interesting events.
- DPD (Derived Physics Data): The AODs are considered as the basis for the analysis. However, in many cases one can easily create even smaller data files which contain only events of interest and for these events, reduce the event data to contain only quantities needed for the end-user analysis. This is done by the process of
  - Slimming: Keep only data used for analysis, in practice, remove whole containers which will not be used.
  - Skimming: Keep only events that pass certain selection criteria like a minimum number of leptons.
  - Thinning: For selected events, store only physics objects that satisfy certain criteria, for example only high quality leptons.

The DPDs span from very general ones, often provided by physics working groups, which are typically called primary or secondary DPDs (D1PD, D2PD) to very specialized ones. These are tailored by single users or small groups for their final analysis with a narrow physics scope. The format can be a POOL ROOT file or flat ntuples (D3PD). The size of D3PDs can be significantly reduced, even fit on a laptop, and the files are capable of a ROOT [90] analysis independent of the large Athena [84] software framework.

The hierarchical structure of the computing model takes into account the varying capacity of the different sites and optimizes the use of our resources. The powerful CERN Computing Center provides both a high number of CPUs and the CERN Castor Mass Storage System. It constitutes the highest, or one could say the most central level, called Tier-0. Its role is to receive the RAW data, copy it to the Castor Storage System tape for archiving and distribute a full second copy over the 11 Tier-1s for storage and subsequent

---

<sup>3</sup>A *run* is defined as a longer period of data taking with the same trigger selection on the same fill in the accelerator.

reprocessing. A first-pass calibration and alignment is run within 24 hours and a first-pass reconstruction within 48 hours. The Reconstruction outputs (ESDs, AODs and TAGs) are distributed to the Tier-1s.

The Tier-1s are based on large computing centers, with many smaller Tier-2 facilities attached to them building up a computing cloud. Tier-1 systems are responsible for keeping a fraction of the RAW data and run more time consuming calibration and alignment procedures, as well as reconstruction with improved calibration, alignment and algorithms, also called reprocessing. Resulting data files are to be distributed to the Tier-2s where they are available for physics analysis access. A Tier-1 is expected to keep a full current set of ESDs and AODs. Regarding MC data, a full set of AODs should ideally also be available. Individual users should in principle not use the Tier-1 facilities for analysis, however they are open to large-scale production of DPDs and analysis jobs usually organized through working groups.

As long as the volume of collision data is small, the Tier-1s are mostly used for managed MC production, however during data taking this activity will be dedicated to the Tier-2 centers. These will also take part in calibration and alignment when and where appropriate. Physicists will also use the Tier-2 sites for their analysis, an activity which is rather chaotic in nature.

If local resources like for example a pool of desktop machines is Grid-enabled, these will be considered as local Tier-3s. For users it may be both an entrance point to the Grid and a local computing and storage facility.

## 5.2.2 Data replication and distribution

In order to assure safe storage of the raw data as well as an equal access to analysis data for all the members of the collaboration, ATLAS applies the following model for data replication [91]:

- RAW: A complete set of the original data will be stored at the Tier-0. An second complete replica will be distributed among all Tier-1s.
- ESD: The output from the primary reconstruction resides at Tier-0 and are exported to 2 Tier-1s. Subsequent versions of ESDs, produced during reprocessing at Tier-1s, are stored locally and replicated to another Tier-1 to have globally 2 copies on disk.
- AOD: These are the common basic data format for most analyses and complete replicas are intended at each Tier-1. Partial replicas will be distributed to the Tier-2s associated to each Tier-1 and the selection may be according to the interests of the local user communities. The rest will be distributed according to capacity.
- TAG: will be replicated to all Tier-1s, both database and ROOT files, while a subset of the ROOT files will be replicated to Tier-2s, each having at least all TAGs matching the AODs stored in the cloud.

Although having a complete private copy of the relevant analysis data is impossible, most users will have some small samples for analysis development and debugging. This step becomes particularly important knowing that it can avoid wasteful use of shared computing resources.

### 5.2.3 ATLAS Distributed Data Management

The Distributed Data Management system (DDM) [92] plays a central role in the experiment as most of the other systems like the managed production and user analysis depend on its performance.

Expecting an overwhelming amount of data, files are organized hierarchically in so-called datasets or rather in larger dataset containers. Such datasets represent some logical data content, like certain real data runs or MC simulation of a specific process indicated by a meaningful name [93].

The DDM is managed by a software stack called DQ2 (Don Quijote 2) [94] which offers both site services and a simple user client. Each Tier-1 system which can also be referred to as a DQ2 site has VO Box [95] (a dedicated node or system for managing VO related services) and a Local File Catalogue (LFC) [96] allowing a distributed implementation of the services. These are the corner stones of the dataset-based bookkeeping system which keeps track of the dataset replicas and their associated files providing information like the GUIDs (Globally Unique IDentifiers), logical file names (LFNs) and metadata like checksums. Data is transported along some default paths between the Tiers using File Transfer Service (FTS) channels.

The link to the physical storage systems at the sites are through Storage Resource Manager (SRM) [97]. Each DQ2 site points to a SRM area, given by an endpoint like `srms://ndgf...`

From the physicist's user point of view the DQ2 client is in practice the main DDM tool providing cross Grid dataset/file management like searching, listing and transferring.

### 5.2.4 ATLAS managed production

The ATLAS MC production and reprocessing is managed centrally. For MC production physics groups prepare the sample definitions files and specify the needed number of events within certain quotas. Groups often collaborate and share samples as there are common interests in some processes, either as signal or background. The ATLAS appointed production manager defines the simulation tasks in the central ATLAS production data base (ProdDB). All parameters needed for an unambiguous definition of the jobs are stored according to a complex schema. A large simulation can be split into several tasks with unique task identification number, each containing a number of jobs. Tasks are then assigned to the different Grids and picked up by the respective production systems which submit the jobs to the computing resources. The status of particular jobs and information regarding output and performance is continuously reported back to the database and an overall status of the ATLAS production system is monitored by the ATLAS Dashboard [98] and PanDA monitoring system [99].

A more detailed description of the NorduGrid/NDGF specific implementation of the production system components will follow in Section 5.7.3.

### 5.2.5 Distributed Analysis

The amount of real data and related MC simulation that will have to be treated in nearly all kind of analyses is enormous. ATLAS has therefore implemented the "jobs go to data" approach for analysis. Users will have to submit their analysis preparation like DPD production or the actual analysis jobs to the Grid as it is practically impossible

to maintain private data copies. In order to provide simple access to the Grid, ATLAS users may choose between two tools. pAthena is a PanDA [100] pilot based command-line submission tool [101] capable of sending jobs to all three ATLAS Grid infrastructures, while Ganga [102] is a more general computational task management tool. The latter is used by LHCb and ATLAS and aims at giving a homogeneous environment for processing data on a variety of resources spanning from local machines or batch systems to the three Grid flavors available for ATLAS users. Ganga handles both general applications like ROOT [90] as well as experiment specific applications. In case of ATLAS there is support for Athena analysis jobs, small scale production as well as the required data management through DQ2 tools.

The flexible plugin architecture of Ganga allowed for a simple integration of a NorduGrid distributed user analysis interface through a so called backend module. The Ganga software as well as the NorduGrid backend will be described in more detail in Section 5.9.2.

### 5.2.6 Monitoring

The global computing environment of the LHC experiments requires reliable, informative and user friendly monitoring systems. The CERN-based developers team maintain the Dashboard project for LHC experiments where ATLAS has its own service covering a variety of Grid activities [98]. It monitors the processing and data management, both at the Tier-0 and on the world wide production system, user analysis and the online/offline databases.

The PanDA [100] production and distributed analysis system provides also a global monitoring system with a web interface [99].

## 5.3 NorduGrid

The ATLAS distributed computing environment is based on three Grid middleware flavors deployed on resources with different geographical location. The US-Grid uses the Open Science Grid Software Stack, Virtual Data Toolkit (VDT) [85] while the LCG sites in Europe, Canada and Far East use gLite [103, 104].

The NorduGrid (NG) collaboration has developed its own set of tools and services which build up the Advanced Resource Connector middleware - ARC [105]. One year after the project start in 2001, developers provided the first architecture and implementation which allowed the establishment of a testbed. Since August 2002, the middleware has been in continuous operation and development. It was initiated by the High Energy Physics (HEP) community which is still one of the largest user groups, but as it is a general purpose middleware, it is also used within other sciences like bio-informatics, climate research and medicine. Figure 5.1 shows the ARC deployment map and indicated resources available for ATLAS with circles.

NorduGrid is currently working closely with other projects. Most important is the Nordic DataGrid Facility (NDGF) [88], a collaboration between Denmark, Finland, Norway and Sweden managing a production Grid facility that leverages existing national computational resources and Grid infrastructures. For the HEP community, NDGF provides the Tier-1 infrastructure and storage services. Due to this collaboration, ATLAS terminology refers often to the NorduGrid resources as to the NDGF Tier-1 site.

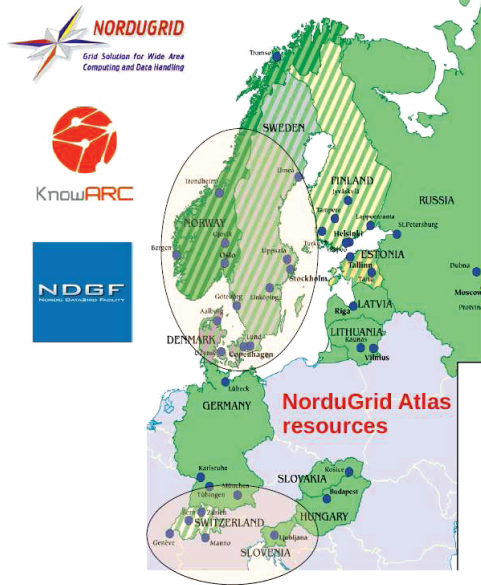


Figure 5.1: Geographical location of the NorduGrid ATLAS resources. Resources available for ATLAS are marked with red circles.

### 5.3.1 ARC architecture

ARC provides a lightweight user client which can be downloaded as a tar file for a number of the most popular operative systems. All parameters defining a job need to be specified using the Extended Resource Specification Language (xRSL) [106]. Such a specification file contains basically a main executable, possibly with arguments, input and output data and their respective locations, as well as parameters describing the job resource requirements like requested CPU and wall time, memory and disk space. A number of community specific software packages are installed on the different clusters by their administrators and made available as so-called run time environments which need to be specified as one of the job parameters. These job requirements are then put into the brokering algorithm which searches for matching resources.

ARC aims at providing a system which is stable by design as it avoids centralized services for vital tasks like data management or job brokering. The architecture involves only a small number of mandatory components [105]:

1. The *Computing System* is implemented as a GridFTP - Grid Manager (GM) pair of services. The Grid Manager is the “heart” of the system and it is instantiated at each computing resource’s front-end. It is a GridFTP gateway to the system and it communicates with the Local Resource Management System (LRMS). All movement of data is managed by the GM which ensures that all input files are in place before the job is submitted to the LRMS as well as stages out the results to external storage. Such a model where all data transfers are handled outside a job minimizes the waste of CPU time as the worker nodes are only used for the actual processing.

2. The *Information System* is the “nervous system” of the Grid. Behind the scenes it consists of a hierarchical distributed database. Information about the services (computing, storage) is stored in the *Local Information System* (LIS) while the *Index Service* maintains a list of available resources.
3. The *Brokering Client* is the “brain” of the Grid system. Given the input from the information system, it aims at powerful resource discovery and optimized matching of the user specified job requirements and available resources.

The work share implemented in this architecture has proven to be both scalable and highly efficient. Jobs are handled according to the so called “push-mode” where the choice of destination is based on a match aiming at an optimized utilization of the resources. It presumes of course an effort to provide correct information input to the brokering algorithm, both from the resources and from users. If users underestimate the job resource requirements like CPU time or memory usage, the jobs will be killed by the LRMS. This strict policy imposes some discipline on the users, while at the same time protects the sites from processes that run out of control. The policy also allows utilization of resources of different capacity. Nodes too weak for a certain kind of jobs can still perfectly well handle other less demanding tasks.

The organizational structure of the NorduGrid collaboration consisting mainly of academic institutions who own and administrate the actual resources made available on the Grid has been a motivation to make ARC a lightweight Open Source software solution. It is made available for a number of operative system (OS) platforms, as well as communicate with different local batch systems. ARC puts emphasis on providing flexible, scalable, non-intrusive and lightweight solutions, both on the server and client side, which allows to connect highly heterogeneous resources.

It is important to point out these features of ARC as they differ from the design of other middlewares. The NorduGrid specific implementation of the ATLAS applications which will be described in following chapters, is strongly dictated by the ARC functionality.

### 5.3.2 Work flow of an ARC-job

As a basis for understanding the implementation of distributed managed production and analysis systems at ARC enabled resources, it is useful to walk through the work-flow of an ARC job. It will clearly illustrate some of the important middleware design features mentioned earlier.

The brokering is done on the client side during submission. Users must specify the requirements of the job like the needed run-time environments (application software), expected CPU-time, wall time and memory usage. The specified input data can be compared with the content of the caches at the sites in order to find if some resources already have a local copy and thus reduce the amount of data movement. The job requirements are matched to the available resources based on the input from the Information System and the job is assigned to a cluster or more precisely, to a queue. Once the job is accepted, it is assigned a unique job identification number and a session directory. The job is further handled by the local GM which is the link to the batch system at the cluster. NorduGrid has a very convenient and informative Grid Monitor [107] which allows the user to know the status of the job at any time. A job which is accepted by a queue, but before the GM starts the preparation of the input data, is in the “ACCEPTED” status (see Figure 5.2 ).



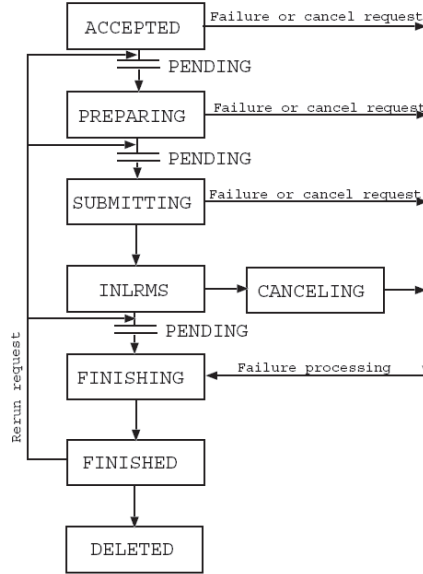


Figure 5.2: ARC job flow and status schema [109]

The fact that all data movement is done by the GM imposes the absolute condition that the users must specify in the job description file (xRSL-file) all inputs and their locations and all outputs and their foreseen final destination. ARC supports a number of transfer protocols. In addition to files in the local system which may be uploaded by the user interface (UI) at submission, input can be assembled by the GM using for example GridFTP, FTP, HTTP(S). Most common for ATLAS is to access input and output data referring to LFC and SRM locations. Before moving to LFC (Fall 2008), NorduGrid used the Globus [108] indexing service Replica Location Service (RLS).

During the process of download of input data, the job is in the “PREPARING” status. Only when all the requested input files are successfully downloaded can the job be submitted to the LRMS (status “SUBMITTING”). If there are no free nodes, jobs are queued with status “INLRMS:Q” which changes to “INLRMS:R” as the actual execution starts.

When the job is finished and waiting for the GM to post-process, it is in status “EXECUTED” and goes into “FINISHING” as the GM starts uploading the output. Then again, only when all requested output files reach their destination, can the job be declared “FINISHED”.

A job can be cancelled by the user at any point in the cycle and it is set to status “KILLED”. If any of the steps fails, the job is set to “FAILED”. If the error occurs during download of the input files or during the stage out of output, the job is marked as re-runnable. This means that by using the resume function the job execution can be restarted from the point where it failed. This feature reduced the waste of data movements and CPU-time in case of a storage or network glitch or downtime. Figure 5.2 gives a schematic overview of the full job cycle.



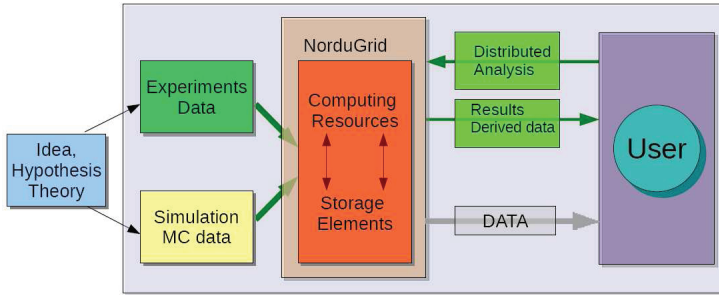


Figure 5.3: NorduGrid - NDGF end-to-end Grid computing environment

## 5.4 NorduGrid/NDGF end-to-end Grid computing environment

The ARC-enabled resources provide a complete end-to-end Grid computing environment for HEP experiments of which ATLAS is the main user, although not the only one. It includes the following components:

- Data Management - covers the full chain from detector to user
- Production System - managed large scale MC production and data processing
- Distributed User Analysis - access via standard user analysis tools

Figure 5.3 gives an overview of the Grid computing environment and workflow. The following sections will give a more detailed description of each of the listed components. A brief introduction describing the data management and NorduGrid resources will be followed by two sections reporting on a Grid logging and ATLAS software distribution project carried out during the first years of my PhD thesis work. Section 5.7 will describe later work related to the production system in NorduGrid/NDGF and its operation, while Section 5.9 will describe the implementation of the distributed analysis.

### 5.4.1 Data Managements - NDGF

The data management and the distributed Nordic Tier-1 is facilitated by the Nordic DataGrid Facility (NDGF). While NDGF does not own any resources, its main activity is in planning, coordination and operation of Nordic Grid storage facilities for major projects among which HEP is one of the largest. In collaboration with the HEP community and with NorduGrid, NDGF is also developing Grid middleware and services. As a result, today there is a fully operational distributed Tier-1 which is integrated in the ATLAS computing system [89].

The NDGF Tier-1 is considered distributed as it does not have a core Tier-1 center and a distinct internal structure of Tier-2s and Tier-3s. This means that all data stored in the system and resources are available as a flat structure, both for the Production System and for the physicists users. The only exception are the dedicated reprocessing queues with access limited to production managers.

| Time      | Application           | Number of clusters | Number of CPUs |
|-----------|-----------------------|--------------------|----------------|
| 2004      | DC2                   | 22                 | 700            |
| Jan. 2006 | Start CSC             | 14                 | 800            |
| Oct. 2006 | DC3 $\rightarrow$ CSC | 16                 | 1500           |
| Oct. 2007 | CSC                   | 15                 | 4000           |
| Nov. 2008 | MC/Data Processing    | 12                 | 9000           |
| Apr. 2009 | MC/Data Processing    | 13                 | 11000          |
| Mar. 2010 | MC/Data Processing    | 14                 | 14000          |

Table 5.1: NorduGrid resources. It is important to note that the resources available for ATLAS users are shared resources and have to handle both production and user analysis.

### 5.4.2 NorduGrid HEP Computing 2004 - 2009

This section gives a short summary of the various ATLAS production campaigns, the development and NorduGrid resource situation.

During Data Challenge 2 (DC2) which took place in 2004, NorduGrid consisted of 22 sites in 7 countries which shared around 3000 CPUs of which 700 were available for DC2 [110]. The resource situation was similar in December 2005 when DC3 started. At that point user analysis was not yet implemented on the ARC connected resources. Towards the end of 2006 the number of CPUs available for ATLAS doubled to  $\sim 1500$  (shared with other users). An initiative was then taken to introduce a NorduGrid backend to the Ganga Distributed analysis tool [102]. This development went on in parallel with the integration of the NorduGrid data management with the overall DQ2 based ATLAS framework.

In April 2007 we could present at the ATLAS Physics Analysis Tools (PAT) meeting in Bergen a GangaNG backend which already supported the main applications and DQ2 datasets as input, although not yet as user output. The collection of sites on which ATLAS software was deployed had by then grown to 20, providing  $\sim 2500$  shared CPUs.

Throughout 2007 and 2008 both the NorduGrid MC production and the GangaNG backend were considerable contributors to the ongoing Computing System Commissioning (CSC) exercise.

In November 2008 the number of shared CPUs available for ATLAS increased to around 9000 provided by 13 sites. The data management at the NDGF Tier-1 was fully integrated with the ATLAS DDM, both within the production system and user analysis.

Parallel to the growing number of CPUs, the storage capacity has grown from 36 TB in 2004 to approximately 830 TB in 2010 (ATLAS related storage space). As of today (March 2010) the number of sites available for ATLAS activities is still around 14. However, several of these have extended significantly their capacity and there are currently up to 14, 000 shared CPUs available. Table 5.1 summarizes the CPU and site development.

During these years, the ARC based end-to-end ATLAS computing system has gone through a considerable development from prototypes to highly efficient production quality system. Readiness tests and continuous running have proved that it is capable of handling the full computing chain of the experiment.

## 5.5 NorduGrid Logger

ARC provides a very powerful monitoring service [107], however it shows only the current status of the Grid system, while there is also a need for a service which keeps the usage history and allows to extract various statistics.

In 2004 NorduGrid introduced a prototype usage data base as well as a web interface, the so called NGlogger [111]. The idea was to provide information which was complementary to the monitoring service, for example usage and performance history according to different selection criteria like users or user groups, applications or sites. This tool provided a simple access to the usage statistics and patterns which were useful for performance analysis, optimization, debugging and documentation.

The prototype database contained the following information:

- start - start time of the job
- end - end time of the job
- cluster - the Grid identification string of the site where the job was executed
- user - user Grid identification string from which one can extract the user name
- name - name of the job
- failure - a short error message in case of failure
- lrms - the Local Resource Management System (LRMS) used at the executing site
- queue - the name of the queue which accepted the job
- rls - the content of the xRLS file specifying the job which can provide detailed information about the job
- ui - user interface, the machine from where the job was submitted

The logger web interface was based on a standard installation of web server with support for MySQL [112] and PHP [113] which was the language used for coding. The graphics were created using the JpGraph [114] package which is based on the GD library [115].

### 5.5.1 Examples

This section describes some examples of the NGlogger functionalities. The front page gave a general overview job statistics covering the past year and an updated histogram showing job statistics for the past week and four weeks.

The menu offered statistics for the different clusters organized in years. For each cluster one could obtain job statistics as well as error listing.

Figure 5.4 shows elements of the Data Challenge 2 (DC2) project page. It offered a possibility of obtaining statistics for certain type of jobs and datasets during a selected time period. A similar page was implemented for the Computer System Commissioning (CSC) MC production. As the number of datasets increased significantly, it was more appropriate to implement queries for various types of jobs like event generation, simulation, reconstruction or merging.

Another typical usage was to search for statistics and information for a single user or a group of users. NGLogger offered a dedicated user page with simple log-in, listing only jobs run by that particular user. Figure 5.5 shows an example of such a user page.

|    | Cluster                                    | Number of jobs | Efficiency [%] |
|----|--|----------------|----------------|
| 1  | <a href="#">dc4.uio.no</a>                 | 3              | 100.0          |
| 2  | <a href="#">lxsrv9.lrz-muenchen.de</a>     | 554            | 99.3           |
| 3  | <a href="#">grid.uio.no</a>                | 5010           | 99.0           |
| 4  | <a href="#">atlas.hpc.unimelb.edu.au</a>   | 697            | 94.1           |
| 5  | <a href="#">hypatia.uio.no</a>             | 4425           | 91.1           |
| 6  | <a href="#">bluesmoke.nsc.liu.se</a>       | 5726           | 88.7           |
| 7  | <a href="#">benedict.aau.dk</a>            | 10583          | 85.1           |
| 8  | <a href="#">sq-access.pdc.kth.se</a>       | 11279          | 81.6           |
| 9  | <a href="#">lheppc10.unibe.ch</a>          | 3268           | 80.3           |
| 10 | <a href="#">brenta.ijs.si</a>              | 16783          | 79.7           |
| 11 | <a href="#">haGRID.it.uu.se</a>            | 5830           | 79.0           |
| 12 | <a href="#">farm.hep.lu.se</a>             | 2124           | 78.2           |
| 13 | <a href="#">inGRID.hpc2n.umu.se</a>        | 14890          | 77.2           |
| 14 | <a href="#">fe10.cdsc.sdu.dk</a>           | 2774           | 77.0           |
| 15 | <a href="#">fire.li.uib.no</a>             | 1893           | 76.7           |
| 16 | <a href="#">genGhis.hpc.unimelb.edu.au</a> | 180            | 76.7           |
| 17 | <a href="#">morpheus.cdgc.dk</a>           | 1589           | 70.4           |
| 18 | <a href="#">sigrid.lunarc.lu.se</a>        | 9523           | 67.1           |
| 19 | <a href="#">atlas.fzk.de</a>               | 12618          | 57.6           |
| 20 | <a href="#">grid01.unige.ch</a>            | 418            | 52.2           |
| 21 | <a href="#">charm.hpc.unimelb.edu.au</a>   | 408            | 45.1           |
| 22 | <a href="#">lsf.nbi.dk</a>                 | 578            | 37.7           |
| 23 | <a href="#">seth.hpc2n.umu.se</a>          | 4              | 25.0           |
| 24 | <a href="#">grid.fi.uib.no</a>             | 16             | 0              |
|    | Average efficiency                         |                | 71.6           |

(a) Cluster efficiency

Specify time:  
 From (MMDD): 0615 Year 2004 To (MMDD): 1231 Year 2004 (def)  
 ATLAS DC2 has been running from 05.10.2004 to the end of December 2004. Some few jobs have  
 Please select a dates in this period for more details.

Select datasets

☐ All (default)

☐ All eventgenerations ☐ All simulations ☐ All digitization

☐ evgen.A1\_z\_ee ☐ simul.A1\_z\_ee ☐ digit.A1\_z\_ee ☐ lumi10.A1\_z\_ee

☐ evgen.B4 Jets17 ☐ simul.B4 Jets17 ☐ digit.A4

☐ evgen.A6\_djet500 ☐ simul.A6\_djet500 ☐ digit.A6\_djet500

☐ evgen.A4\_Wmin ☐ simul.A4\_Wmin ☐ simul.A4\_Wplus

☐ digit.A4\_WplusMU ☐ digit.A4\_WplusTAU

☐ digit.A4\_WminMU ☐ digit.A4\_WminTAU

☐ evgen.0001 ☐ digit.P

☐ simul.H1\_130\_4f ☐ simul.H2\_180\_4f ☐ simul.H3\_120\_gamgam

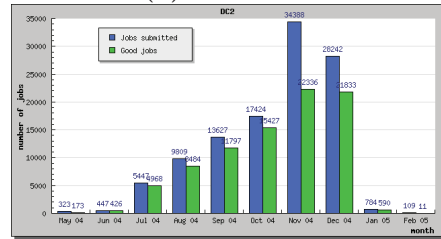
☐ digit.H1\_130\_4f ☐ digit.H2\_180\_4f ☐ digit.H3\_120\_gamgam

☐ evgen.A5\_zJJ ☐ simul.E ☐ digit.E

☐ evgen.A8\_qcd

Display Reset

(b) DC2 search



(c) Job statistics, DC2

|   |        |        |
|---|--------|--------|
| Number of jobs allowed to run             | 111173 |        |
| Good jobs                                 | 86372  | 77.7 % |
| Failed in files download (pre-processing) | 11693  | 10.5 % |
| Job exit code is 1                        | 6849   | 6.2 %  |
| Job exit code is 2                        | 2746   | 2.5 %  |
| Job submission to LRMS failed             | 1279   | 1.2 %  |
| Job exit code is 271                      | 894    | 0.8 %  |
| Job exit code is 127                      | 456    | 0.4 %  |
| Job exit code is -1                       | 196    | 0.2 %  |

(d) Error list

Figure 5.4: NGLogger DC2 overview. Example of a “project page”.

One could query according to a combination of parameters like a time period, certain cluster(s) with additional possibility of a wild card query matching some string pattern in the job name. For the selected jobs one could obtain job efficiency histogram, run time distribution, as well as a more detailed listing for the particular jobs like error messages or the submitted xRSL file which could be useful for debugging. The lists were displayed in a separate window and allowed the user to sort the output, for example according to job name or submission date. This detailed listing was also available for the project pages (DC2, CSC).

One of the problems related to error listing in a highly heterogeneous environment was that the error codes and messages were not coherent. Attempts to list failures according to error codes gave confusing results as the same code could have different meaning at different sites. A more unambiguous failure reporting system based on either a more uniform set of error codes and messages or reliable mapping could improve the situation.

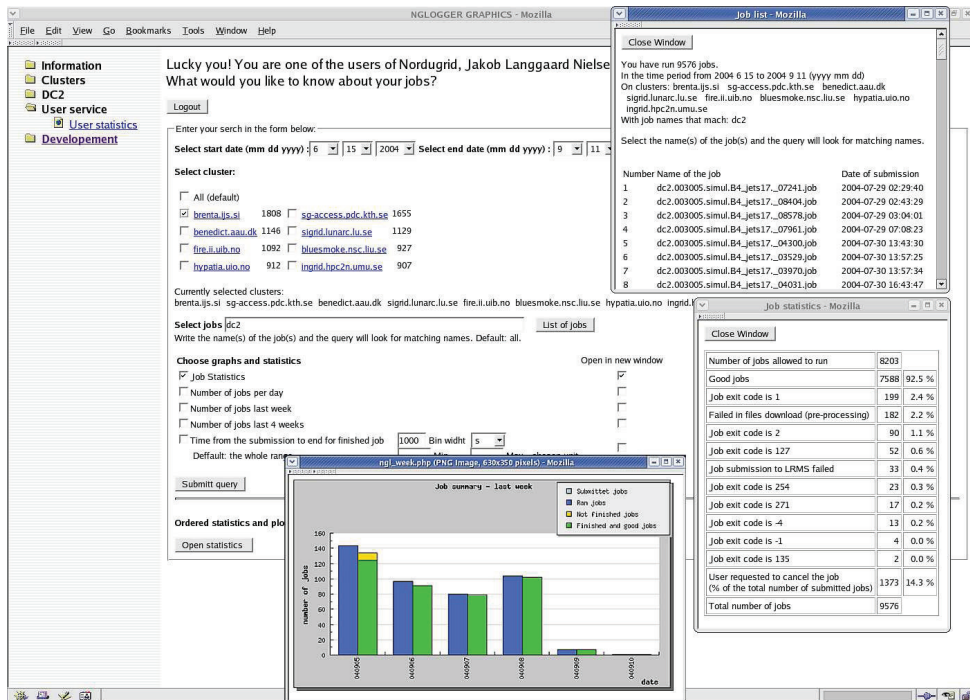


Figure 5.5: NGLogger example of user statistics showing job and error listing, as well as a job statistics graph.

## 5.5.2 Conclusions

The NGlogger prototype proved to be useful and informative. It was particularly used during the early phase of the CSC production before ATLAS had implemented its own global production system monitoring services.

The simple data base design reached a clear limitation as the number of jobs increased and the query time became too long. A more advanced NorduGrid Logger framework and Usage Record has later been developed [116] however, the web interface has not been adapted to the new data base implementation and has been phased out.

## 5.6 RPM-distributed ATLAS software

In the NorduGrid computing architecture, resources connected by the ARC middleware are in most cases not owned or administrated by the HEP communities and the choice of hardware and operative system is in hands of the owner institutions. Compared to the CERN Linux distributions, owner institutions often choose different platforms or more recent releases which in many cases are not binary compatible with the official ATLAS software distribution. During the ATLAS DC2 in 2004 and until approximately summer 2006 there was therefore a need for several specific native releases of the ATLAS software in order to be able to take advantage of these resources [117]. The often non-CERN-related system administrators requested an as light as possible and easy to install production

software kit.

For as long as it was needed, the solution was to provide a nearly Linux-platform independent binary distribution kit of the ATLAS production software. The distribution was based on RPM (RedHat Package Management) [118], a powerful and widely used open-source package management tool.

In addition, the source RPM-files were produced facilitating building procedures on new platforms.

### 5.6.1 The paths of ATLAS

The basis for outside-CERN ATLAS software building is to set up an environment which through a number of locally adjusted environment variables mimic the CERN software environment. The local computing setup which does not allow users any root privileges requires also that all installations are done outside the system directories.

The installation location of each external package is specified with an environment variable of a certain name related to the package name.

As an example, the installation directory of the “gsl-package” is pointed to by the variable `GSLBASE`. In case of system-like packages as “gcc” or “mysql” which are already available through the local system, the related environment variables point to these locations, eg. `GCC_DIR` is set to `/usr`. All HEP-specific packages and ATLAS-specific releases of standard packages, like “python”, are assumed to be installed in the directory specified by the variable `ATLAS_ROOT`. Their respective environment variables point to `ATLAS_ROOT`. A full list of environment variables that need to be set for the building procedure can be found in Appendix A.1.

In order to use RPM as a non privileged user one needs to establish a build directory (`rpmbuild`) in a place where the user has full access and there create the usual sub directories used by RPM `RPMS`, `SRPMS`, `SPECS`, `BUILD` and `SOURCES`. The RPM tools are then configured by letting the `_topdir` point to this location instead of the system RPM area. RPM keeps track of the installed software in a database and an empty `rpm-database` must be initiated in the new setup. To avoid unresolved dependencies this database needs to be filled with information about the software already installed on the system. This is done with the help of a virtual “rpm” that does not contain any actual software, only the system software records.

### 5.6.2 The ATLAS software

The full ATLAS software installation includes a long list of software packages. These have to be built and installed in a certain order due to the many dependencies. A more detailed description of the build procedure and dependencies can be found in Appendix A.2 and A.3, while Table A.1 lists versions and locations of the packages used in ATLAS release 11.0.42.

Once the external packages are installed, the ATLAS software itself could finally be built. The process starts by checking out the entire source code for a release using Configuration Management Tool `cmt` [119,120]. For release 11.0.42 which was one of the last to be built using this framework, the release consisted of more than 1100 packages and more than 630 MB of source code.

Having the CERN-like link structure in place, the building started by setting key locations and `cmt`-variables:

```
$SITEROOT
ATLAS_EXTERNAL=$SITEROOT/atlas/offline/external
LCGexternal=$SITEROOT/sw/lcg/external
LCGreleases=$SITEROOT/sw/lcg/app/releases
$CMTCONFIG
$CMTPATH=<atlas-location>:<Gaudi-location>:<LCGCMT-location>
$CMTINSTALLAREA
```

In order to obtain the correct paths to the external software as defined in the `LCG.Interfaces` package one has to run `cmt` configuration in `LCGCMT` :

```
cd LCGCMT/LCGCMT-<version>/LCG_Release/cmt
cmt config
source setup.sh
```

After the setup followed the main build command (here with the release 11.0.42 as example):

```
cd AtlasRelease/AtlasRelease-11-00-42/cmt
cmt config
source setup.sh
cmt broadcast "cmt config; make"
```

Given the large number of packages and the variety of programming languages used (gcc, fortran90, java, python) there were always packages which appeared to be hard or even impossible to compile. If these were not essential to the foreseen application, mainly Monte Carlo Production, they were enclosed in the `cmt -exclude` option.

The whole build procedure took approximately 11 hours using a Pentium IV machine with 1.0 GB memory.

The first step after finishing the compilation was to install all relevant parts of the software in a temporary location. This basically meant collecting the output from the ATLAS installation located in the `InstallArea` directory, corresponding to the `$CMTINSTALLAREA`, and packing it into four RPM packages.

The ATLAS installation area is organized according to following directory and RPM structure:

```
InstallArea/i686-rhel34-gcc323-opt/bin → atlas-bin-<release>.i386.rpm
InstallArea/i686-rhel34-gcc323-opt/lib → atlas-lib-<release>.i386.rpm

InstallArea/include → atlas-include-<release>.i386.rpm

InstallArea/jobOptions
InstallArea/python
InstallArea/share
InstallArea/XML → atlas-share-<release>.i386.rpm
```



The main application of these RPM distributions was a managed production which only required a binary distribution. However, as the idea of distributed user analysis started to take shape, there appeared a need for compilation of user packages on the Grid in order to avoid incompatibilities. In contrast to the managed production which only used installed precompiled packages, user analysis is based on private code. There are two possibilities. The first one is to provide analysis binaries and direct jobs only to clusters with a compatible system, basically meaning same `gcc` compiler. The second is to provide the source code of the analysis and run the `cmt` compilation on the node and thus obtain compatible libraries. This requires the full structure of Gaudi, `LCGCMT` and the Atlas source code directory structure containing the respective `requirements` files. The source code itself is not needed as the `requirements` files define through `cmt` the correct environment which is defined to point to the installed binary installation. There was therefore an additional rpm package:

```
atlas-requirements-<release>-1.i386.rpm
```

which provides this file structure, the `requirements` files, some required scripts and the `project.cmt` files.

### 5.6.3 Installation and validation

The installation of the complete release is taken care of by one single script. It gives a choice between an installation where the earlier mentioned four key environment variables (`ATLAS_ROOT`, `G4INSTALL`, `ROOTSYS` and `CERN`, see Appendix A.1) are predefined or a fully automatic installation in which these variables are given conventional values, starting from the location where the script is run. The installation of the requirements RPM necessary for compilation is optional.

The command for an automatic installation (option `-a` for automatic, `-p` for predefined environment variables) is:

```
> chmod o+x installatlasrpms-<platform>.sh
> ./installatlasrpms-<platform>.sh -a
```

The installation contains the following steps:

- Make and/or check if the required directories are created (see list in Appendix A.1)
- Download rpm's
- Create and initiate an rpm-database
- Install packages relocating them to correct directories
- Run `cmt.env-RELEASE.sh` which use `cmt` to explore the requirements of the release based on the choice of `$CMTSITE`, and create links to external software
- Run `SealPluginRefresh`
- Check if the most important rpms are installed and that the Seal Plugins are cached



- Generate a minimum run time environment script.

In order to reproduce the full software environment the installation kit provides an environment configuration script, `cmt_env-RELEASE.sh`, which is run after the installation of the requirements RPM. It requires a handful of key environment variables: `SITEROOT`, `ATLAS_ROOT`, `ROOTSYS`, `G4INSTALL`, `CERN`, `CERN_LEVEL`, `RELEASE`, `CMTVERSION` and `CMTCONFIG`. Most of them are automatically set by the installation script since they are related to the installation directory. `CMTSITE` is set to `STANDALONE`. `cmt` is then used to extract the version and directories of Gaudi and LCGCMT and the locations are added to the `CMTPATH`. This is done by using `cmt show macro_value`. A link structure pointing to the external packages is then created based on the information provided by `cmt`. Before running the `SealPluginRefresh` the script checks for existing cache and removes it. The `SealPluginRefresh` is then run on the entire release from scratch.

The whole ATLAS rpm release (11.0.42) counts 2.7 GB. Additional 1.1 GB is needed for the rpms during the installation, but these can later be removed.

Running the whole operation on one of the lxplus machines took about 1 hour, where half of the time was spent on downloading the rpms from the http-site. Installation on a desktop with 1 Mbit connection takes approximately 10 min.

### 5.6.4 KitValidation

ATLAS has developed a test suite which validates some crucial functionalities required for reliable performance. The tests are run against central installation at CERN as well as a validation of the installation on remote sites. The production team in NorduGrid has developed a validation script which provides the necessary environment, runs the KitValidation and collects the validation output in a dedicated directory.

Pacman [121], the official ATLAS software distribution tool organizes the software in a different way compared to the RPM distribution. The environment required by the KitValidation had therefore to be translated to the local setup with the help of a minimal structure of directories and links pointing to the local software installation and by defining some key environment variables like `$SITEROOT`. A uniform setup is an important basis for reliable tests and comparisons. Next, the script downloads the KitValidation and the JobTransforms proper to the ATLAS release to be tested and runs the validation. An example of a KitValidation output is shown in Appendix A.4.

The ATLAS run time environments installed on the NorduGrid clusters have been tested by running the KitValidation as a standard Grid job with the validation script as the executable.

### 5.6.5 Conclusions and outlook

RPM has proven to be a powerful tool for distributing a native build of the ATLAS software. This solution of the compatibility problem allowed the HEP community to take advantage of considerable resources during the different Monte Carlo production challenges from 2004 until approximately the summer 2006.

When the CERN computing system was upgraded to Scientific Linux based on RedHat Enterprise it became more compatible with most of the NorduGrid sites which then could start using Pacman distributions of the centrally compiled software. The compilation

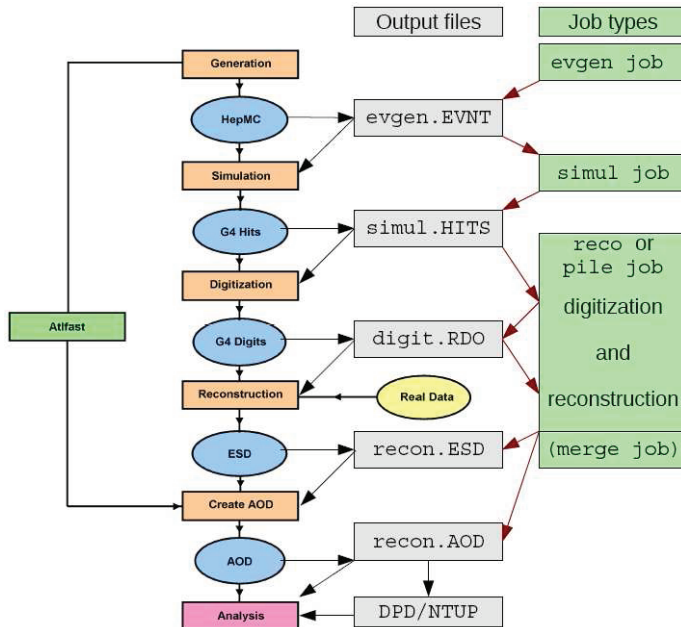


Figure 5.6: MC managed production work flow.

effort in Oslo became obsolete around the ATLAS release 11.0.42. It was thus not updated to the current project organization of the packages.

Compilation of the ATLAS software outside CERN, on a different platform and with different compiler versions, often newer than CERN-defaults, often revealed problems and was a contribution to adaptation of the code for upcoming CERN Linux upgrades. Thanks to the strict requirement of consistent validation results, the effort may also be considered as a contribution to the quality assurance for the ATLAS software and computing infrastructure.

## 5.7 ATLAS Managed production System

The Monte Carlo simulation through the official production system starts with a physics group defining a process and providing a well validated job option file with all process parameter and if needed, generator input and filters. A reasonable number of events must also be specified and fitted within the overall quotas. In the next step, the simulation is defined in the ATLAS production data base (ProdDB) according to a certain schema as a chain of task.

### 5.7.1 Monte Carlo simulation chain

The whole production chain is displayed graphically in Figure 5.6. The description below is valid for ATLAS release 14 and the 2008 MC simulation. This workflow is a result of several years of continuous running and development. While the steps have stayed basically the same, the organization has changed over time towards a higher degree of

automatization. The actual jobs for each step in the chain are defined in so-called job transform Python [122] scripts. These are more complex objects than normal job options as they manage the configuration of skeleton job options based on simple command line arguments, run the actual Athena job and produce job reports.

The production chain starts by generating the physical process using one of the common Monte Carlo generators. ATLAS is using a number of generators in order to provide the best possible simulation of the different processes. The four vector output from the event generation is stored using the common interface of HepMC [123]. These *event generation* jobs are usually short, providing small `evgen.EVNT` output files.

The simulation of the propagation of particles through the ATLAS detector is based on the Geant4 toolkit [124]. It is used to build the full ATLAS geometry and simulate the interaction of particles with matter calculating the response of the active parts of the detector, as well as simulate the inefficiencies introduced by dead material like support structures or cracks. The output is in the form of *hits* carrying information like position of the interaction, “signal” deposits in the detector part involved in the interaction. This step is performed in *simulation jobs* providing output files labeled `simul.HITS`.

In real collisions one expects several interactions at each beam-crossing leading to typically a signal event with several superimposed minimum-bias events (23, at design luminosity of  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ ). In addition, the cavern itself will be filled with particles originating from the beam halo and beam-gas collisions which also have to be taken into account in a simulation as they represent a serious background. The pile-up simulation adds to the process of interest hits from a specified number of minimum-bias events as well as hits generated by the cavern background.

The simulation files (`simul.HITS`), either directly from simulation or from pile-up merging, are then put into digitization jobs in which the various energy deposits are converted into detector responses, “digits”, which correspond typically to voltages or times on preamplifier outputs. This process simulates the detector specific features of the charge collection, including electronics noise, and channel-dependent variations in detector response. This output is then processed by a simulation of Read Out Drivers (RODs). The final output are the Raw Data Objects (RDO) which are in the same format as those containing real data. In the experiment, RDOs will provide “byte-streams” which will be converted to `pool.root` RDO files.

From this point RDOs, whether from simulation or from the experiment, are processed by the offline reconstruction software in so-called reconstruction jobs. The output contains now physics objects like tracks, particle candidates or energy deposits, and is written to detailed ESDs and the more analysis friendly AODs.

The simulation step is the most CPU-time consuming (1000 s/event) and a single job handles only of the order of 50 events leading to an output of about 2MB/event (ca. 100 MB per file). While the reconstruction jobs are considerably shorter (20-40 s/event). A “one-to-one” task chain is therefore unpractical for several reasons. The small number of events per simulation job results in a large number of small simulation files which is a problem for the Distributed Data Management (DDM) system. The corresponding short reconstruction jobs (order of 15-20 minutes) are, on the other hand, a problem for the executors as the run-time is too short compared to the job processing time which includes preparation, submission and post-processing. A single executor, processing jobs continuously can handle only a limited number of jobs per day. Therefore, a large number of short jobs in the system makes the utilization of the resources inefficient as it creates

a pileup of finished jobs combined with drained queues.

It has therefore been decided (from around release 15) to merge the simulation files (`simul.HITS`) into “JumboHITS” of 1000 events (ca. 2 GB per file). Even with merged HITS input, the resulting AODs are still considered too small and numerous and have to be combined in a merged AOD containing of the order of 10,000 events (ca. 2 GB per file).

The digitization and reconstruction steps may be done in a combined reconstruction job without storing the RDO files and only using them as temporary files within the particular jobs. Avoiding storing the RDO files saves a considerable amount of disk space (2 MB/event). In a similar way, the simulation and digitization may also be performed as combined job which provide RDO files while the HITS files are temporary.

For less detailed studies there is a possibility to run a simplified simulation called ATLFAST [125]. It takes the event generation files as input and provides AODs as output.

### 5.7.2 ATLAS Production System

The starting point of the production system is the Production Database (ProdDB) which contains definitions of all jobs and information regarding their processing. Each step in the simulation chain is defined as a task which consist of a number of jobs. The output of a task is collected in the basic data unit of the DDM, in a so-called dataset. For processes which require a large number of events the work load may be distributed over several tasks, often handled by several Tier-1s. The output is then gathered under a common dataset which takes the form of a dataset container. Users can access the files via DQ2 data management tools. Metadata describing the processing parameters like ATLAS geometry or release, as well as cross sections and filter efficiencies can be obtained via the ATLAS Metadata Interface (AMI) [126,127].

The ProdDB job definition and bookkeeping is organized in a schema consisting of three tables: ETASK, EJOBDEF and EJOBEXE. These are shortly described in Appendix A.5.

The access to the ProdDB is restricted, but the relevant information may be accessed through the monitoring system. Until fall 2007 very useful overviews, status reports and statistics were provided by the ATLAS Production System Monitoring [128]. Later it was integrated in the ATLAS dashboard [98]. Users outside the production team may obtain information about different datasets and tasks, such as their status, through the AMI database or the PanDA monitoring system.

The software tool running the production is a supervisor-executor pair. The supervisor is common for all Grid flavors, while the executor is the Grid specific implementation of the required functionalities and an interface to the computing resources. Figure 5.7 shows a schematic overview of the ATLAS production system indicating the various components.

The Eowyn supervisor is the steering unit for running production. It is implemented as a Python class which uses several helper objects like the `LoggingMgr` and `LogDBProxyOracle` for logging. The `ProdDBProxyOracle` object provides methods for interacting with the ProdDB, like fetching jobs waiting for processing and keeping the job records updated. Eowyn is thus the only component which directly interacts with the ProdDB. Once the jobs are finished successfully, Eowyn calls for DQ2 registration with the help of the `DMSDQ2_PROD` class which is the production system DQ2 interface.

The main stages of the Eowyn workflow are [129]:

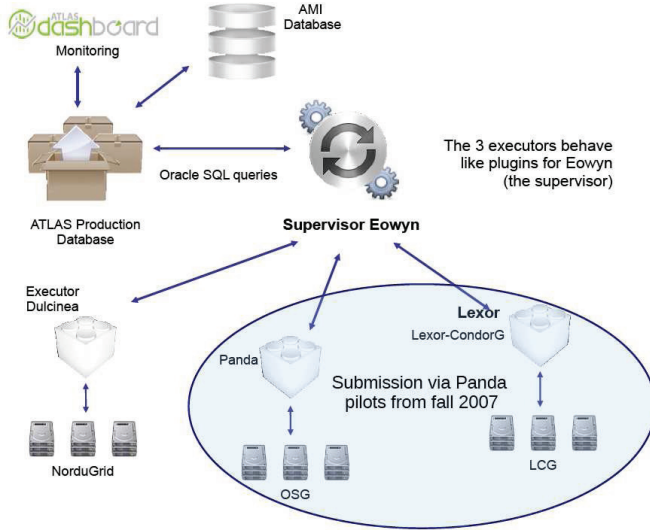


Figure 5.7: ATLAS Production System. Schematic overview.

- Select available jobs from the ProdDB  
This is done according to an order clause. The ordering is according to the priority of the jobs and the task number which assures that older tasks are prioritized in order to promote their finalization. The maximum number of jobs to be submitted per cycle is set in the executor. With rare exceptions, it has been kept at 100.
- Pass the jobs to the executor for preparation and submission  
Here the jobs are passed to the executor as a Python object containing all the information necessary for a complete job definition which is then translated into a Grid specific job description, xRSL in case of Nordugrid. Most of the information is from the EJOBDEF's JOBPARS field (see Table A.3). JOBLOGS and JOBOUTPUTS-fields are used for a correct specification of the input and output files. This is important for the DDM bookkeeping, and it fits well with the ARC design which requires specification of all input and output and their locations in the xRSL.
- The Grid specific executor then takes care of the preparation and submission of the jobs
- Eowyn is periodically calling a method that checks and reports back the status of the jobs.
- Finished jobs are post-processed:
  - Successfully finished jobs are reported back to the ProdDB, Eowyn calls for DQ2 registration of the files and declares the job as DONE.
  - Otherwise the jobs are registered as FAILED and released for a resubmission.

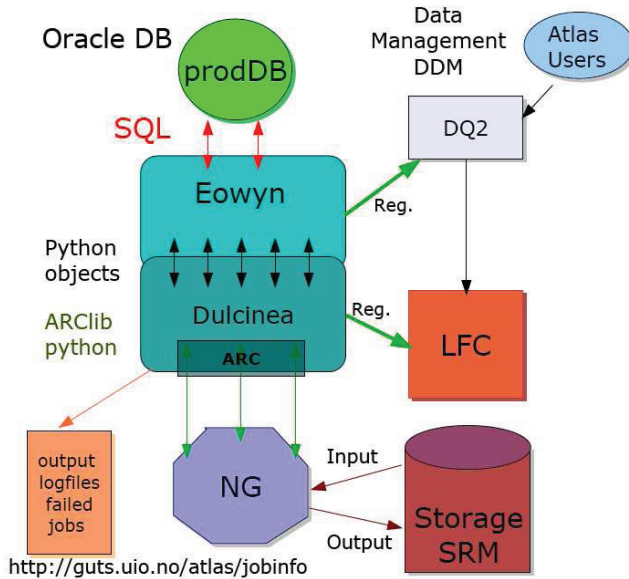


Figure 5.8: NorduGrid production system with the Dulcinea executor and Eowyn supervisor. Jan. 2009 design.

### 5.7.3 Dulcinea

Dulcinea is the NorduGrid specific implementation of the functionalities required by the ATLAS production system, a so-called executor class, whose methods are called by the common ATLAS supervisor Eowyn. All executors follow the same structure allowing Eowyn to be a generic tool for all Grid flavors.

It is the executor's task to parse the job objects from Eowyn, create a Grid specific job definition and submit the jobs. In Dulcinea this means creating an xRSL file with a correct configuration. The executor class provides also methods for checking the status and post-processing of finished jobs, as well as clean and kill methods.

Dulcinea is also written in Python [122] and based on the ARClib Python binding which provides a very flexible and powerful middleware toolbox [130]. Figure 5.8 shows the final design (Jan. 2009) of the NorduGrid production system based on a supervisor/executor pair.

The main methods will now be described in more detail following the systems workflow. Some comments regarding the development are also included.

#### getNumJobs

The routine asks for the number of jobs that the executor is ready to receive and which Grid/Tier-1 it is sending them to. In Dulcinea the implementation of this is very simple as it only returns back the tier name "NDGF" and a fixed number. Given the fast submission and the ability of the system to distribute the jobs, we have found 100 to be a reasonable number. Complementary quantities are the Eowyn parameters `maxActive` and `targetMaxActive` which imposes a maximum number of jobs per executor. If the limit is reached, Eowyn simply skips the submission cycle.

This maximum limit is manually set and adjusted according to the resource status as viewed in the NorduGrid monitoring system, where one can also see concurrent activities, like user analysis, the length of the running jobs and the number of queued jobs. Although the interventions are manual, it does not represent a heavy work load as a quick look or a hint from the production team is easily passed to the executor. Thanks to a reliable brokering system, the tuning does not have to be very precise or frequent. On the other hand, this flexibility allows for a maximal exploitation of the resources.

## prepareJobs

Next Eowyn queries the ProdDB for a batch of available jobs and calls the `prepareJobs` method of Dulcinea. Each job definition is parsed and the xRSL description is constructed.

For each of the input files assigned to a job one has to obtain the GUID from DQ2. With the growing amount of MC data, querying the central DQ2 database for each file became a very slow process. In order to speed up the submission, if some of the jobs are from a task which is new to the executor, a local DQ2 cache is populated using dataset bulk query. The cache is created for a whole dataset and subsequent handling of jobs from the same task requires only a quick look-up in the local file. A new DQ2 query is performed only if new files have been added to a dataset.

When creating the xRSL, the Grid jobs are given the same name as the `JOBNAME` in the ProdDB. These give an indication both of the physics process and job type (e.g. `evgen`, `recon`, `simul`) and correspond to the output dataset name which follows the ATLAS nomenclature [93]. This job naming convention is very useful as the job names are displayed in the monitoring system and make it easy to spot problematic tasks.

In ProdDB jobs are defined to be executed using a certain release of the ATLAS software and Production Cache (e.g. `TRANSINFO` field in the `EJOBDEF`, Appendix A.5, Table A.3). The job transforms to be used and the `cmt` configuration are also specified. With some amount of string manipulation, this information is parsed and translated into the xRSL `runtimeenvironment` attribute which specifies the software environment to be set up on the node. The `runtimeenvironment` attribute is an important argument for the brokering algorithm as only clusters providing the required software are taken into account. The release information is also passed as the environment variable (`T_RELEASE`) and together with the job transform, they are added as entries in the executable `arguments`.

Input files prescribed for the job are specified either as single files, `inputLFN`, or as lists `inputLFNlist`, for example:

```
type=LFN metatype=inputLFN value=
<PFileInfo lfn=EVNT.023753._00015.pool.root dataset=mc08.105003.pythia_sdifff.evgen.EVNT.e344_tid023753
PFileInfo>
type=LFNlist metatype=inputLFNlist value=[<PFileInfo
lfn=HITS.028171._[86895,86896,86897,86899,86901,86902,86903,86905,86906,86910].pool.root
dataset=mc08.106020.PythiaWenu_1Lepton.simul.HITS.e352_s462_tid028171 PFileInfo>]
```

After unpacking the input file list short-notation, GUIDs are obtained from the local DQ2 cache, while the LFC URLs using a DQ2 helper function which translated the dataset name into an LFC path (`LFCtools.to_native_lfn`). Before the full DDM integration (finalized summer 2008), GUIDs were obtained from the NorduGrid RLS. The input location was defined to be a RLS-url. At this stage, if an input file is missing the GUID,



the whole job preparation is stopped as this indicates incorrectly registered or even missing file(s) and the job would never succeed. This sanity check was particularly useful in the early days of the production when the system was still under development.

Output file names are also extracted from the job object and assigned an output location url.

Currently this is an SRM-url indicating the space-token and location, for example:

```
srm://srm.ndgf.org;  
spacetoken=ATLASMCDISK;  
overwrite=yes/atlas/disk/atlasmcdisk/mc08/AOD/<dataset>/AOD.042061._28475.pool.root
```

If the job is rerun because of failure, an overwrite mode is enabled. This is necessary as some files from the previous job may have been successfully written to the storage system. The earlier used RLS-based system did not have this feature. Jobs that only partly succeeded in upload or crashed at a point when (some) output files were already stored outside the worker node, would fail in subsequent attempts in upload as it would be blocked by the already existing file. In order to avoid this, for all jobs that were rerun, an attempt was made to remove the foreseen output files. This additional operation, often performed for 2-3 files, slowed down the submission and thus reduced the total number of jobs that could be handled by an executor per day.

Further, the job description is equipped with a number of configuration attributes dealing with for example logging or queue rejection. However, most important are the `cputime` and `memory`. As the numbers provided by the ProdDB were often underestimated, they were replaced by more realistic numbers based on experience. Well tuned values of these attributes are very important. First they are used in the brokering process where the memory requirement may exclude some resources. Once the job is submitted and exceeds the indicated requirement, the job is killed by the Grid Manager.

Finally, the main component of the job, the actual `executable` is set. It is not the athena run command itself but the `NGExecWrapper` [131] shell script containing all commands that are to be executed on the node. On the node the script is executed with the `arguments` string as input. It contains a long list of parameters used in configuration and execution of the actual Athena job (an example of an `arguments` list can be found in the box below).

The `NGExecWrapper` shell script performs following actions:

- The input file names and GUIDs in the `arguments` string are used to create a `PoolFileCatalog.xml` in the session directory.
- The runtime environment and `cmt` setup is completed.  
A combination of the middleware `runtimeenvironment` feature and configuration steps carried out by the `NGExecWrapper` is capable of providing successful sequential setup of two runtime environments within one Grid job.
- The Athena execution command is composed using the required job transform and specified parameters which configure the job.
- The Athena job is executed.



- After the Athena job is finished: check sums of the output files (md5sum and Adler32) are obtained. These are, together with other metadata like the sizes and GUIDs of the output files, summarized in the `OutputFiles.xml` which is downloaded during job post-processing. Such meta data required for the latter file registration can only be obtained from the files themselves and has to be collected before the files are uploaded to the storage location.
- Log files and some selected files from the session directory are gathered in tar files for logging and debugging purposes.

Below follows an example of a xRSL job description for a simulation job.

```
&(middleware >= nordugrid-arc-0.6.5)
(jobname = mc08.105003.pythia_sdifff.simul.e344_s512_tid042832._01723.job)
(executable = NGExecWrapper)

(runtimeenvironment = APPS/HEP/ATLAS-14.2.25.6-I686-SLC4-GCC34-OPT)
(environment = (T_RELEASE 14.2.25))

(arguments = "2" "EVNT.023753._00050.pool.root.2" "0A0C6C88-A250-DD11-9095-00145EFDCA6"
"DBRelease-6.3.1.tar.gz" "744d633f-b0c4-49f8-aae1-f0a1d454cc79"
"2" "HITS.042832._01723.pool.root" "None"
"mc08.105003.pythia_sdifff.simul.HITS.e344_s512_tid042832"
"log.042832._01723.job.log.tgz" "None"
"mc08.105003.pythia_sdifff.simul.log.e344_s512_tid042832" "0" "1"
"preinstalledCache" "APPS/HEP/ATLAS-14.2.25.6-I686-SLC4-GCC34-OPT"
"csc_atlasG4_trf.py"
"EVNT.023753._00050.pool.root.2 HITS.042832._01723.pool.root
50 1100 1723 ATLAS-GEO-02-01- 00 QGSP_BERT jobConfig.TruthModVertexPos.py
DBRelease-6.3.1.tar.gz NONE False ")
(inputfiles =
(NGExecWrapper
lfc://lfc1.ndgf.org//grid/atlas/dq2/user/user.alexread.production/NGExecWrapper9)
(EVNT.023753._00050.pool.root.2
lfc://lfc1.ndgf.org//grid/atlas/dq2/mc08/EVNT/<dataset>/EVNT.023753._00050.pool.root.2)
(DBRelease-6.3.1.tar.gz
lfc://lfc1.ndgf.org//grid/atlas/dq2/ddo/DBRelease/<DB>/DBRelease-6.3.1.tar.gz))

(outputfiles = (metadata.xml "") (OutputFiles.xml "") (wrapper.log ""))
(jobSmallFiles.tgz "") (jobInfo.xml "")
(HITS.042832._01723.pool.root
"srm://srm.ndgf.org;spacetime=ATLASMCDISK/<path>/<dataset>/HITS.042832._01723.pool.root")
(log.042832._01723.job.log.tgz
"srm://srm.ndgf.org;spacetime=ATLASMCDISK/<path>/<dataset>/log.042832._01723.job.log.tgz"))

(stdout = log.042832._01723.job.log) (join = yes) (project=atlas) (queue!=atlas-t1-repro)
(queue!=atlas-t1-reprocessing) (queue!=serial_4gb) (gmlog = gmlog)
(jobreport=https://grid.uio.no:8001/logger) (gridtime=1503) (memory=2000)

<path> for example:
atlas/disk/atlasdisk/mc08/HITS or /atlas/disk/atlasdisk/mc08/log
```

If the creation of the xRSL job description is unsuccessful, the job is reported back to the ProdDB as failed with a proper error code. In the opposite case, the submission process starts by creating a `X = Xrsl(xrsl.sling)` object (defined in ARCLib [130]) and the job description is validated. If passed, a list of queues matching the requirements of the job is created using the `ConstructTargets(queueList, X)` method of ARCLib. The resulting target list is passed as input to the brokering algorithm `PerformStandardBrokering(targetList)`.

| ARC - state  | Eowyn - state                 |
|--|-------------------------------|
| ACCEPTING, ACCEPTED<br>PREPARING, PREPARED<br>SUBMITTING | PENDING<br>PENDING<br>PENDING |
| INLRMS:Q   | SUBMITTED                     |
| INLRMS:X ( $X \neq Q$ )<br>EXECUTED, FINISHING, KILLING  | RUNNING<br>RUNNING            |
| FAILED, KILLED, DELETED                                  | FAILED                        |
| FINISHED   | FINISHED                      |

Table 5.2: Map of ARC job states (see Figure 5.2) to Eowyn states.

Dulcinea allows the production manager to blacklist clusters by specifying them in a dedicated file (`.dulcineareject`) which is consulted during the creation of the target list. Queues hosted by blacklisted clusters are excluded from the target list.

The brokering process returns a target list where the resources are ranked according to their specifications and current capacity. Finally, the jobs are submitted using the `JobSubmission(X, targetlist)` object to the most suitable resource according to the brokering algorithm. A job is successfully submitted once it is accepted by a cluster and obtains a unique Grid job identification string which is the address of the job throughout the processing.

Jobs are usually submitted in batches of 100 and the process is much faster than the update of the information system used by the global brokering. What happens is that a cluster which at the beginning of the submission has capacity to accept a certain number of jobs may receive the whole batch as the update of the queue status in the Grid information system lags behind. This is of course far from optimal as the whole batch may be more than the resource can handle in a timely manner. Therefore, in order to obtain a better distribution of jobs within one submission cycle, a queue list internal to the process is updated to include the jobs that have just been submitted. The next jobs in the batch then use the locally updated broker information as it represents a more correct picture.

### **getStatus**

Eowyn is configured to call for a job status check with certain time intervals. The `getStatusJobsEXE` calls the executor `getStatus` method which using the ARCLib function `GetJobInfo` obtains the status information of jobs provided in an argument list.

The first Grid status of a job after it is submitted is `ACCEPTED` and as the Grid Manager starts to stage in input files the status is set to `PREPARING`. These initial native ARC states are mapped to Eowyn status `PENDING`. If a job has to wait in the queue before execution (Grid status `INLRMS:Q`) the corresponding Eowyn status is `SUBMITTED`. As the status on the Grid changes to `INLRMS:R`, the Eowyn status is set to `RUNNING`. The usually very temporary states `EXECUTED`, `FINISHING`, `KILLING` are still mapped to `RUNNING` while the final ARC states are logically mapped to either `FAILED` or `FINISHED`. The full map of ARC job states to Eowyn states can be found in Table 5.2.

The ARC middleware provides a very useful resume functionality. If a job fails during

staging in or out of the files, it can be resumed from the point in the work flow where it failed. A typical example is a storage glitch or downtime. In case of download glitch one does not waste the data which has already been transferred, completing only the missing files. In upload failures for successful jobs, the output is kept on the node and a retry at a later point when the system is again fully functional can prevent considerable waste of data transfer and CPU time.

For finished jobs, Dulcinea downloads log and metadata files, performs the necessary registration of output files and reports back run details to the ProdDB.

The downloaded `OutputFiles.xml` contains information that can only be obtained from the files themselves like GUID, check sums and size. It is parsed and the files and information about them is registered in the LFC. The same metadata is also registered in DQ2.

In case of failures much effort is put into providing short and informative error messages as well as systematized error codes. One of the challenges is to translate the many different error codes and messages into a common schema.

The ProdDB is also a source of monitoring and performance statistics. The job report contains therefore also timing information, CPU model in order to normalize the CPU time usage and number of events that were processed.

With the full integration of the NorduGrid/NDGF production system and ATLAS DDM, files become available to the users as soon as they are registered. In the original model based on the NorduGrid RLS (Globus Replica Location Server) the whole data management was limited to the internal system and was not automatically accessible to ATLAS users through the common tools like DQ2. Additional registration processes were run and data had in many cases to be replicated to other sites in order to become available. An intermediate step towards integration was to add a LRC (Local Replica Catalog) database in RLS to enable DQ2 to communicate with the NorduGrid file catalog. It became obsolete once the LFC was in place. A crucial step towards the full integration was also the commissioning of the NDGF managed SRM/dCache storage system. Currently, both LFC and SRM are fully integrated in the ARC middleware data management. The data handling in Dulcinea is illustrated in the right column on Figure 5.8.

## Killing and cleaning

At times it happens that a task is being called back and the jobs that are running are requested to be killed (jobs with status TOBEABORTED in ProdDB). For this purpose Dulcinea provides the `killJobs` method.

Jobs that have been successfully post-processed have to be removed from the worker node. The `clearJobsEXE` method of Eowyn calls the Dulcinea method `clearJobs`. For successfully finished jobs it simply cleans the work directory. While in case of failure, the whole work directory with all log files is downloaded to a web accessible location. The ATLAS Dashboard Production System monitoring provides a listing of failed jobs and automatically includes a link to these log files.

After the download, the failed jobs are also removed from the worker node.

## 5.8 Production System operation and development

NorduGrid has been in continuous operation since Summer 2002 when it contributed to the first ATLAS Data Challenge (DC1). The author's involvement in the operation of the ATLAS Production System started in December 2004. The following sections are an attempt to group in threads different operation and development issues as well as summarize them in a time ordered way. Given the complexity of the system and the long and radical developments, it is impossible to include all.

### 5.8.1 Storage and data management

As the CSC production started in December 2005 the storage and data management followed still the model used during DC2 which was entirely based on the ARC middleware and the RLS. The Grid system itself was robust, however the production framework still suffered from some early child diseases. The RLS based system did not allow to overwrite a file if uploaded a second time. Such RLS blocking could occur if the earlier failed attempt managed to upload some files. The other and more common case was a Dulcinea crash during submission cycle. This lead to so-called "orphan jobs" on the Grid, jobs that did not reach the registration in ProdDB and were unknown to the production system once it was restarted. These jobs could run and finish successfully. In the mean time, the restarted executor would resubmit the job with the same parameters including identical output files. If the "orphan jobs" finished before the registered one, it would fail in upload. It would succeed only at the third attempt, thanks to the deletion of output files from the previous runs that was always done before job resubmission. This practise improved the success rate, but was a time consuming part of the job preparation, reducing thus the total daily job throughput. Running the same job three times was of course also an inefficient resource utilization. The current SRM storage provides an overwrite capability.

Originally most NorduGrid Storage Elements protected files via the Grid Access Control Lists (GACL). It implied by default that only the owner of the file had access to them. However, the various steps of the production chain may be done by executors run by different people who then need full access to already existing files. Setting an extended GACL was automatically done during the job post-processing. This, however, sometimes failed and lead to failures in subsequent attempts or steps in the production chain. For some time one had to add GACLs by hand as the errors surfaced. The problem was solved by introducing directory GACLs which defined the access rights to the entire content of the directory.

Production System assigned usually the whole production chain of a physics process to only one Grid so that input files were available within the system. When input from other tiers was required, and before the full integration with the ATLAS DDM, it was necessary to check if all input files were available. If not, they had to be transferred by hand as there was no automatic replication.

Simulation output produced by NorduGrid had on the other hand to be distributed using a semi automatic system. During fall 2006 an SRM-less temporary solution was found. A LRC (Local Replica Catalog) was set up together with a simple SRM endpoint using dedicated "channels" (managed by NDGF) [132].

In March 2007 the NorduGrid production system could switch to an SRM/dCache [133] based storage and by fall, the same year, the storage and data management system of the NDGF Tier-1 was fully integrated into the ATLAS system. The RLS was finally phased

out during fall 2008 and the production system registers automatically the metadata in LFC and DQ2.

The new storage system has not changed the ARC data handling philosophy where the Grid Manager is responsible for the data movement. From a user point of view, the only change is that instead of the RLS-URLs the input files now are specified with an LFC-URL, while the output is destined to an SRM-location with a specified space token.

### 5.8.2 Job Error analysis

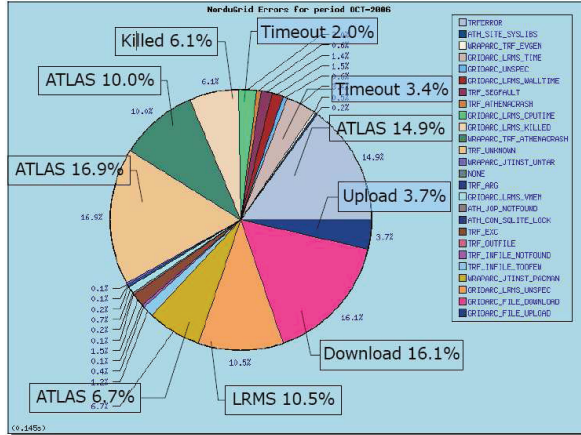
The Production System operates with two types of efficiencies: job efficiency reflecting the job success rate and wall time efficiency reflecting the resource usage. To illustrate the difference, failures early in the job processing lowers the job efficiency, while the impact on the wall time efficiency is small. Long jobs, on the other hand, which fail at a late stage, for example in upload, represent a big CPU-time loss and lower the wall time efficiency. The impact of such failures may be small on the job efficiency if the problem concerns few but long jobs.

For the production in NorduGrid/NGDF, one of the main error sources has been the ATLAS software. Developers are responsible for debugging the code before it enters a release and there are well established testing routines. However, many problems appear only when the code is run on other types of processes than those used in testing, or they appear first in large scale runs. In other cases there are bugs that have simply been missed. Figure 5.9 shows the NorduGrid job errors during October 2006. Although ATLAS related error sources have always been dominant, this period was particularly difficult as the releases 12.0.3 and 12.0.31 suffered from high error rate. The software was responsible for 48.5% of the job errors and 46.9% of wall time errors. The errors in download of input files which occur early caused 16.1% of the job errors while only 0.6% of the wall time failures. The different timeouts (CPU-time, wall-time) caused 6% of the job failures, while the corresponding failed wall time was 15%.

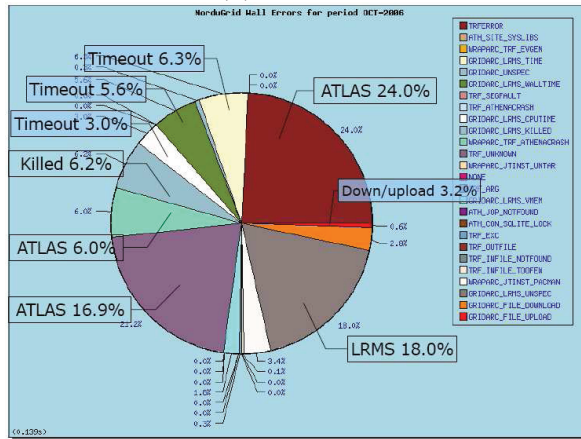
A related error source is the strict limit imposed on the memory usage of a job. If the required amount is exceeded, jobs are killed by the LRMS. These kinds of errors are “expensive” in terms of wasted computing time. Therefore, it has always been important to validate the resource requirements suggested in the task definition. In very many cases the numbers provided by the ProdDB have been highly underestimated. A systematic validation procedure will be described in Section 5.8.3.

Towards the end of 2006, after the first year of the CSC production (processed with ATLAS release 12) we had observed numerous examples of difficult error messages. When reporting the bugs it often appeared that the messages classified as ERROR were actually warnings. On the other hand, errors classified as FATAL did not necessarily stop the process and jobs could continue until they finished or the error caused a segmentation violation. This processing continuation after a serious error caused considerable CPU-time waste and made it difficult to find the real cause of failure. The problem was after some time solved by better error severity distinctions.

Figure 5.10 shows the error break down for September 2007 when the production was run using some of the last releases in the major release 12 and possibly some first of release 13. Again ATLAS software errors are clearly the main error source responsible for ca. 75% of both the job and wall-time errors. The 6% of job failures due to input file download problems are not visible in the wall time chart. The timeout errors accounting for 5% of the job errors correspond to 18% of wall time loss. As mentioned, the main



(a) jobs errors

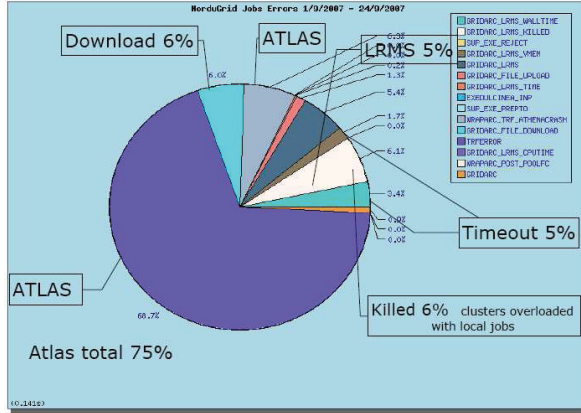


(b) wall time errors

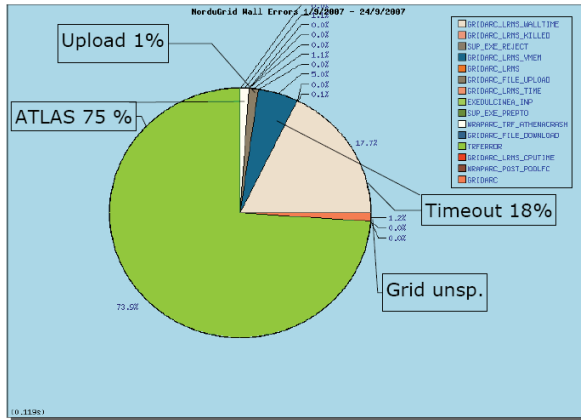
Figure 5.9: NorduGrid Production System errors October 2006.

reason for these is an underestimation of resource requirements. But there has also been peculiarities like certain clusters that systematically use a few percent more memory for certain types of jobs than others. This had to be taken into account when setting the memory attribute in the job description. There has also been tasks with extreme CPU time requirements, sometimes due to the complexity of the events, in other cases jobs entered infinite loops and had to be aborted. The extreme examples are from tasks run with one of the last releases in cycle 14 where jobs were looping and timing out after 2 - 3.5 CPU days.

To complete the story, Figure 5.11 shows the error distribution from November 1st until December 15th 2008 which was a very active production period (200 K jobs, wall time 130 years, job efficiency 94%, wall time efficiency 97%). At this point the production was done using release 14 of the ATLAS software. The fraction of failures caused by the ATLAS software is significantly reduced compared to previous statistics. The LRMS errors (GRIDARC\_LRMS) account for more than 50% of the job errors, however this



(a) jobs errors



(b) wall time errors

Figure 5.10: NorduGrid Production System errors September 2007.

error source is hardly visible in the wall time error chart (accounts for only 0.67% of the wall time error). A typical cause of such errors has been hardware failures like bus errors. A few problematic nodes can potentially fail a large number of jobs in a short time. Looking at the failed wall time, the timeouts are by far the dominant error source.

From these statistics one can clearly see the advantage of the Grid Manager being responsible for the data movement. The job failure listing shows an almost equal number of failures in download of input and stage out of output files (ca. 6%). The upload failures are not shown independently among the job errors in Figure 5.11 chart (a), but they are included in the category “others”. When looking at the wall time errors in Figure 5.11 chart (b), download failures are negligible, while the upload failures are responsible for ca. 6% of the inefficiency.

### 5.8.3 Operation issues and tools

A large part of the production shifter duties has been to follow up the failed jobs, report errors, whether related to ATLAS software or to sites. In the early days of production



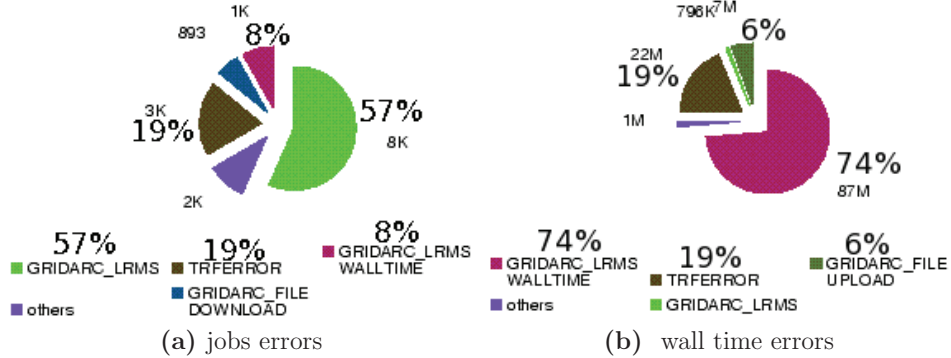


Figure 5.11: NorduGrid Production System errors November, December 2008.

this required much time and manual work. There has been continuous development, both towards a more stable system, in terms of middleware and site performance, as well as towards more automatized and powerful production operation tools. This section will describe some of the important issues and improvements from the start of the CSC production in November 2005 until March 2009.

When jobs in the production system fail, an error code and acronym is registered in the database. If it fails two or three times <sup>4</sup> with the same error, Eowyn sets the job in the AUTOABORTED state. Such jobs need to be investigated in order to find the reason for the repeating failures. In NorduGrid we also tried to let the job run on resources running different platforms in case this could have an effect.

As shown in the previous section, failures related to the ATLAS software have been one of the main factors reducing both the job and wall-time efficiency. A large part of the daily production operation has therefore been to study the log files and submit bug reports to the Savannah [134] bug tracking system. Once a problem was understood and/or reported, job could be set to ABORTED which is considered as one of the valid final job states.

There is also another mechanism that prevents a job from being resubmitted many times. When created, a job is granted a certain number of attempts, usually three (MAX-ATTEMPT field in the EJOBDEF table). The LASTATTEMPT field in the same table counts the attempts and when the LASTATTEMPT reaches MAXATTEMPT a job will no longer be picked up for submission. For such jobs one needs to check the error messages and possibly the log files. As NorduGrid consists of a variety of systems, it often happens that the same type of failure leads to different error messages or codes making an automatic decision whether a job is to be aborted impossible. However, the mapping of the Grid messages to the common Production System error schema has been continuously improved. Other common reasons why jobs got stuck in MAXATTEMPT were system problems like failure in input file download, LRMS or other “trivial” reasons.

One of the daily operation tasks was therefore to list all jobs stuck in AUTOABORT and MAXATTEMPT. This was done using the `maxattempt.py` script [131] which for each job queried the ProdDB and provided the error code and short error message for all attempts from the EJOBEXE table. The same script could then be used to add attempts to jobs that should be retried, and released jobs that were in AUTOABORT, but should

<sup>4</sup>The number of repeated failures of the same type is configurable. This is a reasonable default choice.



still be given a chance. Jobs that failed repeatedly due to a known problem and did not have a chance to succeed were set in status `ABORTED` using the `abortJobs.py` script [131]. The latter can abort jobs in a task taking different kinds of arguments like certain pattern of error messages or acronyms or job identification numbers.

As already mentioned, the resource requirements in the Proddb have often proved to be underestimated. In NorduGrid we developed therefore a so-called “scout-job” strategy. For each new task, we release only a small number of jobs, typically 10, and the rest of the jobs was submitted only after these finished successfully. The “scout-jobs” have often also revealed jobs that are simply not possible to run due to a bug and the whole task was redefined. The release of scout jobs and the rest of the jobs once the task was validated was managed by the (`submitScoutJobs.py`, `submitValidatedJobs.py`, `submitWaiting.py`) scripts available in the ATLAS CVS repository [131]. This “damage control policy” has been very efficient and was implemented in the central production system.

During the years of Dulcinea operation we had two periods of frequent crashes related to the ARC middleware. In Spring 2006 the ARC client suffered from severe memory leakages leading to a growing memory consumption. After a couple of submission cycles, each of 100 jobs, the executor had to be restarted in order to clear the memory, otherwise it crashed. A failure during submission and before the registration of submitted jobs in the Proddb, led to so-called “orphan jobs” as described in detail earlier in Section 5.8.1. The error was fixed by the ARC middleware developers.

In the beginning of 2008 the client suffered for a long time from a bug that lead to segmentation violation during submission. The executors run on new and fast computers were most vulnerable to this problem. An ad-hock solution was to run the executor under Valgrind [135], a memory and threading debugging tool, which slowed down the process and allowed it to run smoothly.

While waiting for client errors to be fixed by the middleware developers, Dulcinea was equipped with internal bookkeeping of submitted jobs. The Grid identification numbers of the jobs were written to a file during submission. The record was deleted once the jobs were registered in Proddb. In case of a crash before this step, a list of unregistered jobs was available and the restarts began with the killing of these “orphan jobs”.

Dulcinea has through its operation time been through continuous development. One of the driving factors was the development of NorduGrid and the ARC middleware. As mentioned earlier, job distribution had to be improved by an internal brokering. New “exceptions” had to be added as new problems came up. On the other hand, the development of the middleware simplified many operational issues. Several new features like the SRM file overwrite capability, which allowed omitting the file cleanup before job resubmission, and the local DQ2 cache, shortened the job preparation time and increased the daily throughput.

One curious ATLAS software related problem lead to jobs producing large log files with the size in the GB range, reaching more than 10 GB. When such jobs failed, the log file download could block the system, as well as quickly fill up the web-available disc area allocated for the log files.

The `NGExecWrapper` script, which is the executable of every production job has also been through considerable development, and has always been up to date with respect to the various new requirements of the Production System.

The ATLAS production is a very challenging use case and over the years it has revealed

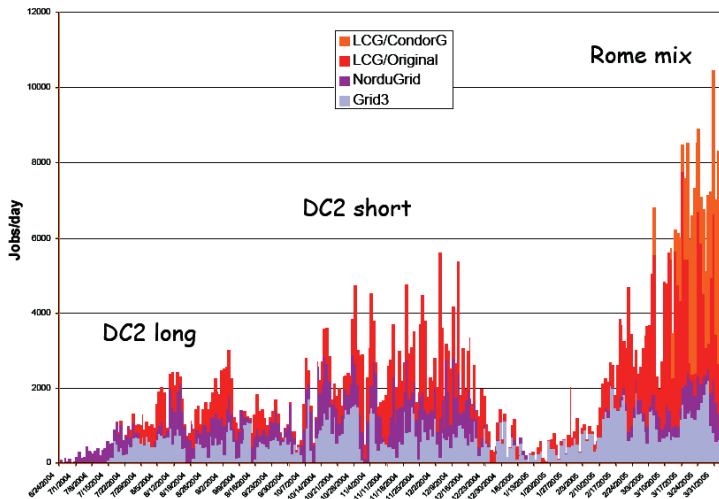


Figure 5.12: Number of jobs during DC2 and Rome production in the three Grids from end of June 2004 until the end of February 2005. NorduGrid jobs are shown in purple.

errors and weak points both in the middleware and in the site configurations. The fast submission and the high number of jobs required improvements of the brokering algorithm as well as a correct site information in the information system used during the brokering. The development and the resulting improved performance has been achieved thanks to excellent collaboration with the middleware developers and the site administrators.

#### 5.8.4 Resource and performance statistics

Since Data Challenge 1 (DC1) which started summer 2002 and was one of the first major applications, NorduGrid has been in continuous 24/7 operation and has developed to be one of the largest production Grids. The increasing number of jobs, clusters, storages, sites, run-time environments and users has so far not caused any serious scalability problems.

ATLAS Data Challenge 2 (DC2) (2004-2005) was the next large scale MC production of 15 M fully simulated events using Geant4, with the new C++ algorithms, and the Athena framework. NorduGrid participated in DC2 with approximately 700 CPUs, distributed over 22 sites. Thanks to high efficiency, NorduGrid delivered 33% of the total production which was comparable with the shares of the two other Grids which participated with nearly three times more resources [110].

During spring 2005 NorduGrid participated in the short term production of approximately 5M events intended for the ATLAS physics workshop in Rome (summer 2005) and contributed 5%. Figure 5.12 shows the number of jobs per day during the DC2 production in 2004 and the Rome production, early 2005. The Rome production was run from Oslo in collaboration with Uppsala and Lund. The job management was handled by the Windmill supervisor and in case of NorduGrid, by the early implementation of the Dulcinea executor (pre CSC design) [84].

In November 2005 the Oslo Experimental Particle Physics Group [136,137] took over

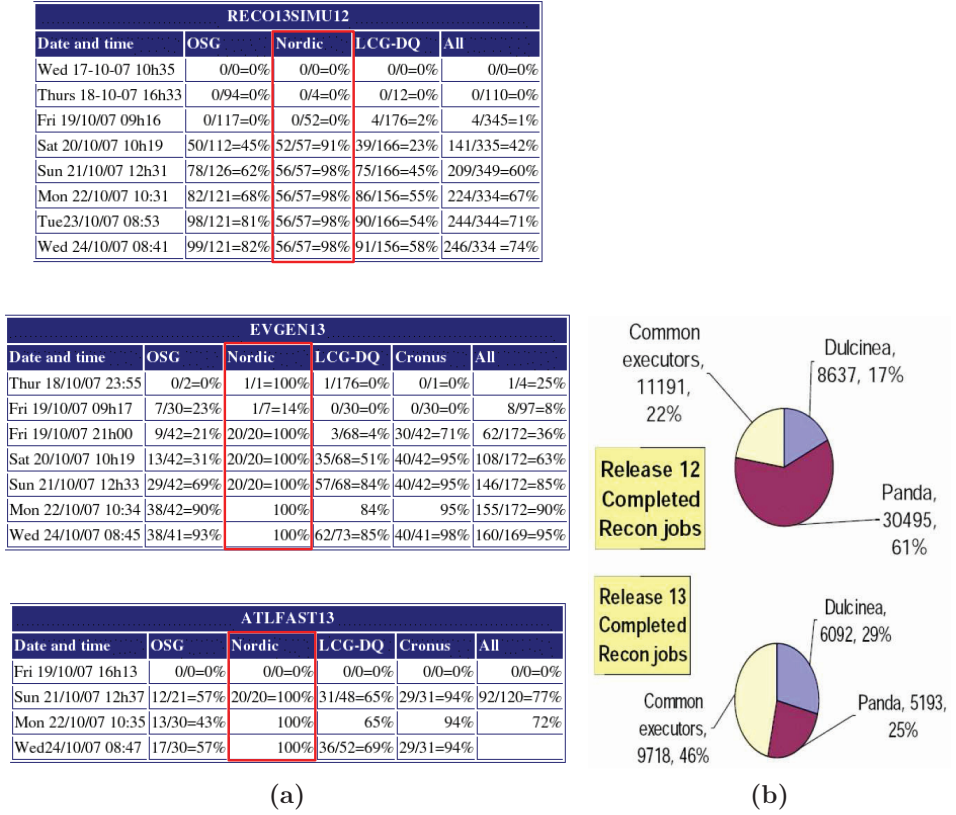


Figure 5.13: Progress time-line and completion rate for high priority validation tasks (a). Overview by Andreu Pacheco during the Oct. 2007 ATLAS Software Week. Full chain completion rate (b). From Kaushik Dee during the Oct. 2007 ATLAS Software Week

the development of the Dulcinea executor and started the preparations for the upcoming Computing System Commissioning production (CSC) which was intended to be a large scale test of the whole system, from simulation to analysis. During the first year, Dulcinea production system was still in its infancy. As described in the previous section, it suffered from many failures, but went also through continuous development. The results of this effort became visible in 2007 when one could observe a significantly higher number of concurrent jobs and improved efficiency. By 2008 the system reached stable running, although there were always challenging issues. When the number of ATLAS software related errors was low, it could keep the production going almost unattended with high efficiency. During the CSC production the NorduGrid resources were integrated in the Worldwide LHC Computing Grid (WLCG) as one of the Tier-1 centers.

Figure 5.13 shows some highlights from the 2007 production. Panel (a) shows the progress time-line and completion rate for high priority validation tasks. The system responded rapidly to the introduction of high-priority tasks and had a high finishing rate within a short time. Panel (b) shows the overview of the production volume on the three

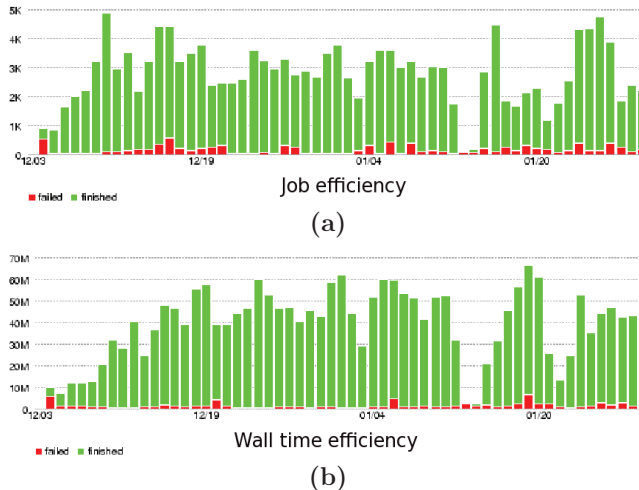


Figure 5.14: NDGF Tier-1 production December/January 2007/2008 was run almost unattended with high throughput and efficiency.

Grids. The NDGF Tier-1 delivered 17-29% of the production full-chain (data which were useful for physicists) with 5-6% of the ATLAS computing and storage resources.

Figure 5.14 shows the NDGF performance during December/January 2007/2008 production which was run almost unattended. The throughput was stable and high resulting in a job efficiency of 94.5% and wall time efficiency of 96.7%.

2008 was in general a year of stable high efficiency production. During the first part of the year there were frequent segmentation violations during submission on fast computers. However, these were properly treated by the “orphan job” mechanism which is described in Section 5.8.3). An automatic immediate restart of the executor and a killing of all “orphan job” from the interrupted submit cycle, limited the resource waste to a minimum. Starting the 2008 production, there was still a need for substantial manual intervention such as task management, failed/stuck job follow-up and replication of special datasets (minimum bias, pile-up). During the year, the earlier described “scout job” procedure was implemented in the central task management and became one of the central “damage control mechanisms”.

In 2008 the NDGF Tier-1 completed the effort to comply with “common standards” of distributed data management and became fully integrated in the ATLAS Production System using the same tools (dCache, SRM, LFC metadata catalog instead of RLS). At the same time, the NDGF Tier-1 is unique in WLCG with its distributed design enabled by the ARC middleware as well as data movement entirely taken care of by the middleware. During this year of development and consolidation it has proven to be highly efficient, reliable and stable.

Figure 5.15 shows the job statistics for period 03.15 - 07.01 in 2008. The upper plot (a) shows the number of successful jobs per day while the lower plot (b) shows the successful wall time per day. The light blue contribution of the NDGF Tier-1 is noticeable and stable. The first period from mid March until the beginning of May 2008 was a production of “real data like” Monte Carlo simulation for the Final Dress Rehearsal 2 (FDR2) exercise and represents a longer stable run period.

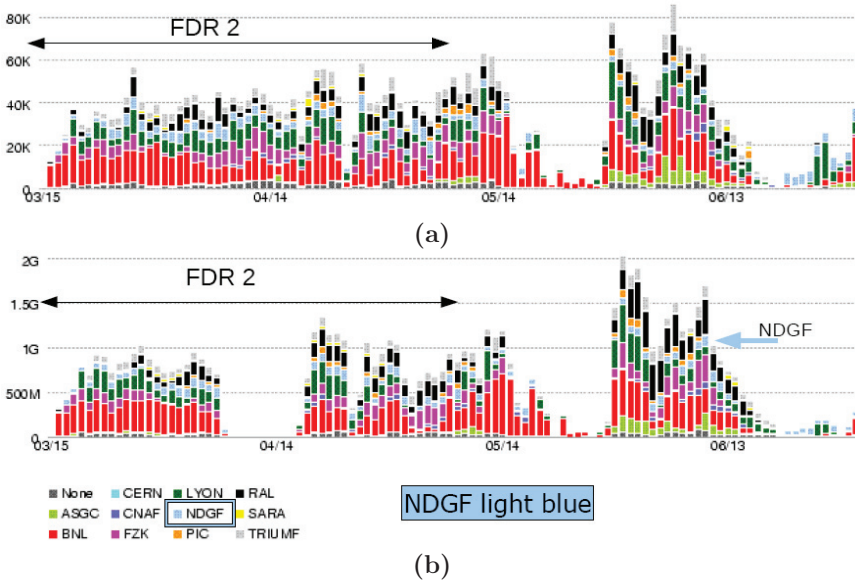


Figure 5.15: NDGF 03.15 – 07.01 2008. Number of successful jobs (a), successful wall time (b). The production from mid March until the beginning of May was the so called Final Dress Rehearsal 2 (FDR 2). The production at NDGF shows a characteristically reliable performance.

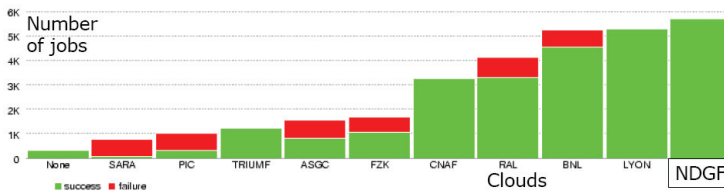


Figure 5.16: Number of jobs processed in 12-hours on the 02.07.2009. NDGF Tier-1 outperforms the other clouds.

In addition to the reliable performance, NDGF Tier-1 has distinguished itself in comparison with the other tiers. There were frequent 12-hour periods when the NDGF Tier-1, considered a so-called “5%” cloud in terms of resources, contributes nearly at the level of BNL, a “25%” cloud. There were even 12-hour periods when NDGF had the highest job throughput an example being 02.07.2009 shown in Figure 5.16, with nearly 100% efficiency.

To summarize, in 2008 the system reached a certain level of maturity. Figure 5.17 shows the job and wall time efficiency during the entire year. The upper table shows part of the cloud-wise statistics, while the table in the lower left corner shows the over all Grid-wise efficiency. The NDGF Tier-1 had a job efficiency of 92.5% and a wall time efficiency of 96.9% which ranked it the best performing tier/Grid.

| cloud     | success  | failure | success (walltime) | failure (walltime) | efficiency | efficiency (walltime) |
|-----------|----------|---------|--------------------|--------------------|------------|-----------------------|
| BNL       | 3476946  | 840232  | 64820033035        | 13205481113        | 80.5%      | 83.1%                 |
| LYON      | 1635723  | 631781  | 28750146192        | 9680811659         | 72.1%      | 74.8%                 |
| RAL       | 1181084  | 464903  | 28481607352        | 9829990071         | 71.8%      | 74.3%                 |
| FZK       | 1067185  | 283652  | 19750650330        | 3092718421         | 79%        | 86.5%                 |
| NDGF      | 1168915  | 95668   | 20287398960        | 639731940          | 92.4%      | 96.9%                 |
| TRIUMF    | 1064815  | 120766  | 17858372744        | 2135305309         | 89.8%      | 89.3%                 |
| grid      | success  | failure | success (walltime) | failure (walltime) | efficiency | efficiency (walltime) |
| LCG       | 7163500  | 2515341 | 133238000098       | 32305323699        | 74%        | 80.5%                 |
| OSG       | 3476946  | 840232  | 64820033035        | 13205481113        | 80.5%      | 83.1%                 |
| Nordugrid | 1143668  | 93333   | 19735311900        | 630388200          | 92.5%      | 96.9%                 |
| None      | 315379   | 332853  | 6149893115         | 1404526366         | 48.7%      | 81.4%                 |
| total     | 12099493 | 3781759 | 2.23943238148e+11  | 4.7545719378e+10   | 76.2%      | 82.5%                 |

Figure 5.17: Efficiency 01.01 – 12.31.2008 compared to some of the other clouds (upper table) and as a summary for the three Grids (lower left corner).

### 5.8.5 Outlook

As writer of this thesis my involvement in the development stopped as the NDGF production system became integrated in the global system. The emphasis moved to DDM which was outside the scope of my project. My later contribution was as a production “shifter”. In the beginning of 2009 this task was taken over by a central production team organized in global 24/7 coverage shifts. This is related to the integration of NDGF Tier-1 in the global production system. Currently (February 2010) the actual job handling does no longer follow the supervisor/executor (Eowyn/Dulcinea) model, but is moved to the pilot job based PanDA [100] framework. On the NDGF side, jobs are handled by the ARC Control Tower (aCT) which integrates parts of the Dulcinea code in a new framework.

## 5.9 ATLAS distributed analysis - Ganga

The distributed Grid computing and data management have been established as the working environment for all LHC experiments. This is the case not only for large scale Monte Carlo or data processing, but also for end user analysis. A full set of real data and simulation corresponding to one year of running is of the order of several Terra Bytes which is beyond the capacity of most local resources. Any upcoming analysis will have to access data accumulated on common distributed storage infrastructures and follow the “jobs go to data” strategy. While the distributed computing systems are very complex in nature, users, on the other hand, will always require simple and uniform access to the resources. Such an interface has to hide as much as possible the technical details, be quick and easy to learn and give a consistent interface to all kind of available resources, from local batch systems to the world-wide Grid. Switching from one type of resource to another should not require any reconfiguration of the application. A typical use case is to develop the analysis on a desktop machine or similar using a small set of data. Then one could test it on some local batch system before a full run on the Grid.

In order to meet these requirements ATLAS provides Ganga [102], a distributed analysis user interface and job management tool. It supports the main ATLAS analysis framework, Athena, as well as integrates the DDM tool DQ2. The basic idea of Ganga is to provide an interface where a user can configure a job once and run anywhere.

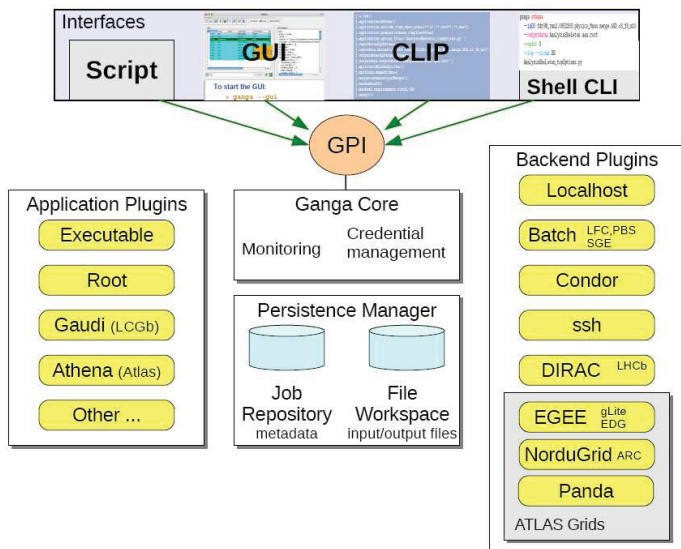


Figure 5.18: Ganga architecture. Ganga offers a variety of user interfaces. The software has a flexible modular plugin design.

### 5.9.1 Ganga architecture

The Ganga software has an object oriented design programmed in Python [122]. Users may access the different functionalities through three types of interfaces: Linux Command Line Interface (CLI), interactive Python command line interface or file based Python script, or a Graphical User Interface (GUI). These cover different needs and use cases. A new user may be more comfortable with a GUI or a quick and simple command line submission. While an advanced user would appreciate the rich configuration possibilities or the flexibility of the Python command line or scripts. Regular users often develop powerful private scripts which automatize generation of frequently run jobs.

These interfaces are built on top of the Ganga Public Interface (GPI) which gives access to the Ganga core components. The Ganga architecture consists of a Core module which takes care of common tasks like monitoring and credential management. All jobs are stored in a job repository and input/output files may be stored in a dedicated Workspace. These functionalities are organized by the Persistence Manager. The architecture of Ganga integrates different components as plugins. Ganga supports a number of different types of jobs through Application plugins. Some are experiment-independent like the generic executable or ROOT, others are experiment-related like Gaudi for LHCb and Athena for ATLAS. Similarly, the different computing resources are interfaced through backend plugins. This modular architecture as sketched in Figure 5.18 makes it very simple to add new applications and backends. Ganga was therefore a natural choice for a distributed analysis frontend to the ARC enabled resources.

The basic entity in Ganga is a job which is defined according to the following schema:

- Application – what to run

Ganga supports the experiment-independent applications “Executable” and ROOT,



as well as a number of experiment-specific ones. In case of ATLAS the most important are Athena and AthenaMC.

- Backend – where to run

The user may easily switch between different resources by changing the backend. This module implements the technical part of job preparation as well as all methods necessary for executing and interacting with the jobs. Ganga offers backends to batch systems (PBS, Condor, SGE, LSF) and to distributed Grid enabled resources. For ATLAS in particular, this means LCG, OSG through the PanDA backend and NorduGrid.

- Input Dataset – Data read by application

The input sandbox is a generic mechanism for file transfer which can be used by any application. There are also more specific “inputdata” methods like ATLASLocalDataset for local files. The DQ2Dataset is based on the build in DQ2 client and manages entire datasets and container datasets which are the basic data units in ATLAS. NorduGrid offers also the application-independent NGInputData which may contain a list of files stored on NorduGrid.

- Output Dataset – Data written by application

As for input there is an application-independent output sandbox. The ATLAS specific ATLASOutputDataset can be used to store output data on a local file system, while the DQ2OutputDataset stores output on the Grid as a user dataset which is available through DQ2.

- Splitter – Rule for dividing into subjobs

The DQ2JobSplitter is an ATLAS related method. It is powerful and practical as the job splitting is based on the data file locations (job-to-data). It also handles container datasets in a correct way. DQ2JobSplitter imposes also a maximum limit on the amount of input data per job. The data files in a given dataset are evenly distributed over the sub jobs. If the size of input data share per job exceeds 5 GB, the DQ2JobSplitter adds more sub jobs.

- Merger – Rule for combining output

The implementation of the Ganga job object is based on a set of basic component classes which inherit from interface classes as viewed in Figure 5.19. The plugin classes define their job attributes through a common schema specifying visibility of an attribute, syntax, as well as a short documentation. Thanks to the schema it is possible to offer a similar user interface for the same type of plugins, like for example for the different Grid backends. At the same time one can add component-specific features.

## 5.9.2 GangaNG - the NorduGrid backend

This section contains a more detailed and technical description of the implementation of the NorduGrid backend GangaNG.

The GangaNG package contains both the implementation of the backend class and the application run-time handlers (in release 5.4.5 for Root and Athena). In addition, the ARC stand-alone client is provided as one of the external packages. It is a lightweight



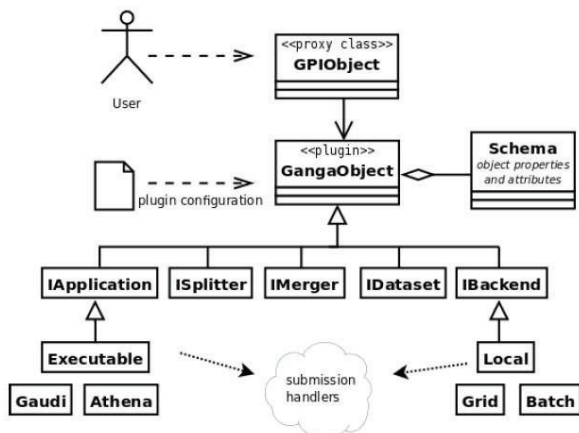


Figure 5.19: The structure of the Ganga component classes and their abstract interfaces. [102]

package, version 0.6.5 sums up to only 23MB, and allows users to submit jobs from any network connected machine without requiring the CERN afs.

Ganga is based on a number of component classes each inheriting from an interface class (see Figure 5.19). For the NorduGrid backend the central class is the **NG**(**IBackend**) inheriting from the **IBackend** interface class. In line with all other plugin classes, it defines its attributes through a schema which specifies their type (read, write-only, internal), visibility, associated user-convenience, syntax shortcuts and short documentation. An additional requirement schema (**NGRequirements** class) provides the Grid specific parameters of a job like CPU-time, wall-time and memory requirements.

The Ganga design aims at making the backend classes application-independent. This is done by distinguishing between the backend object and the so-called submission or run-time handlers which define the application according to the appropriate configuration schema (for NG defined in the **NGJobConfig**(**StandardJobConfig**) class). This schema contains all information needed in order to run the application on a specific backend. The actual preparation of the job done by the backend consists merely of translating the application configuration into a Grid job specification (xRSL) and submitting it to the Grid. All interactions with the Grid-job is through the **Grid** helper class which provides all necessary middleware functionalities.

The monitoring of the jobs is handled by the Ganga Core with input from the backend methods which interact with the Grid jobs via the middleware. Finished jobs are post-processed by the backend which downloads the log files and any output specified to be downloaded. In case of ATLAS applications using DQ2 dataset as output method, the results are uploaded to a Grid storage by the Grid Manager at the site where the job is run. The backend downloads metadata related to the output and performs the required registrations in order to make the files available through DQ2.

Failed jobs may be resumed if they fail during the stage in or stage out phase. This is the same functionality as implemented in the managed production. It is particularly useful in case of successful jobs which fail for example due to some temporary storage problem. With a resume of the file upload one can still finish the job and avoid wasting

time and resources. If resume is not possible, one can resubmit the whole job or only the failed subjobs.

### **GangaNG – Technical issues, implementation, solutions**

The plugin design of Ganga is flexible and makes it easy to introduce new backends and applications. At the same time, there are clear guidelines for which functionalities and methods need to be provided in order to fit within the machinery. A different type of boundary conditions is the functionality of the middleware. This section focuses on some technical issues, implementations and solutions specific to the NorduGrid environment.

One of the requirements has been that the usage of the GangaNG backend should be independent of any centrally installed middleware, for example on the CERN afs. This has been possible thanks to the lightweight ARC stand-alone client which is distributed as a part of the Ganga software. However, the ARC client contains several slightly modified Globus [108] libraries which overlap with those used in the LCG middleware and there has been several problems related to setting environment variables like `LD_LIBRARY_PATH` and thus mixing the middleware environments. The Grid methods in GangaNG could therefore not be based on the very useful ARC Python binding, like in the NorduGrid production system executor. Instead it uses the command line interface via a shell command wrapper which inserts the installation path of the client binaries. The location of the ARC client is always known through the environment variable `ARC_LOCATION` which is set during the configuration of the GangaNG package. In this way one can avoid setting environment variables like `PATH` or `LD_LIBRARY_PATH`.

The necessity of distributed analysis becomes obvious as one starts dealing with large data samples. It is then an advantage to split the job in several subjobs with a smaller number of input files and obtain results in a timely manner. In such cases users also require quick submission of a large number of jobs, often solved through so-called bulk submission. This feature is supported by the ARC middleware. Technically this only requires that the subjob xRSLs are collected in one file and the brokering is done only once. Since the brokering process involves a status request to a number of sites, one can risk to wait for several of them to time out if they are not responding. GangaNG implemented therefore the possibility to set a shorter time-out. A limit of for example 20 sec is usually enough for well functioning sites. However, the users may set it even shorter if they are interested only in a smaller range of fast responding sites.

Additional speed up is achieved through a number of other improvements. For example, the sandbox is created only once. In case of Athena jobs, steps which are common for all jobs are done only once, like the preparation of the user area and the group area, if requested. If using a DQ2 dataset, the DQ2 look-up is performed once. If the user or group area or the sandbox tar ball is larger than a configurable limit, it is uploaded to a temporary Grid storage, only once, and defined as one of the input files to be used by all jobs. When a job is finished (or failed) these files are automatically cleaned.

As already mentioned, ARC design delegates all data movement to the Grid Manager on the site front-end and it is therefore necessary to specify all input and output files in the xRSL job description. The GangaAtlas built-in DQ2 functionality gives a list of files and GUIDs for a given input dataset. Using an LFC tool (`to_native_lfn`), we parse the dataset name and combine with the SRM endpoint from TiersOfAtlas in order to construct the SRM path of the file. Similarly, if DQ2 dataset is chosen as output method, the SRM location of the output files is constructed by parsing the user dataset name. The analysis

output is stored in the ATLASCRATCHDISK (earlier ATLASUSERDISK) space token.

The actual executable of the Athena application is the wrapper script (`athena-ng.sh`). It parses a string of arguments that contain all input file names and their GUIDs. Based on these it creates the `PoolFileCatalog.xml` which is expected by Athena and a list of input files for the Athena EventSelector service. When the Athena job is finished, the wrapper script generates an xml-file (`OutputFiles.xml`) containing information about the output files that can only be obtained from the files themselves like size, GUID and check sums (md5sum and Adler-32 checksum). This `OutputFiles.xml` file is downloaded by Ganga and the entries used in DQ2 and LFC registration of the files. The `athena-ng.sh` script gathers also time stamps used in tests and monitoring of the processing.

For a long time jobs frequently failed because the MC DQ2 datasets distributed by the ATLAS DDM were missing physical replicas of files that were registered in the dataset<sup>5</sup>. In GangaNG users could switch on a check of the availability of the files listed as members of a DQ2 dataset. Files which were not replicated to NDGF Tier-1 were removed from the input list and the analysis was run successfully on the available subset. Unfortunately the check takes a couple of seconds per file and can sum up to a considerable delay for a large dataset. However, during the CSC analysis, it ensured efficient analysis despite of missing files. The check became obsolete after the introduction of the DQ2Splitter which does a similar check.

### 5.9.3 Installation and configuration

Ganga is a free open-source software which is publicly available through an installation script. The installation starts by downloading the script and making it executable. The help function explains the different options. The example below shows the installation of release 5.4.5 in a specified directory and with the ATLAS-related external packages.

Ganga installation:

```
> wget http://cern.ch/ganga/download/ganga-install
> chmod u+x ganga-install
> ./ganga-install -help
> ./ganga-install --prefix=/scratch/ganga --extern=GangaAtlas,GangaNG,GangaPanda 5.4.5
> /scratch/ganga/install/5.4.5/bin/ganga
```

Starting Ganga for the first time one will be asked if one wants to generate a config file `/.gangarc` in the home directory. The `gangarc` file contains a long list of configuration parameters related to various modules and applications. The list below focuses only on the few which are necessary for running ATLAS jobs using the NG backend.

A typical ATLAS user needs to do some small changes in the configuration file. The file contains different sections indicated by the module name in square brackets. The Configuration section contains the `RUNTIME_PATH` variable which specifies which packages to load when starting Ganga. A useful option is to redirect the log file directory and output to some convenient place by setting `gangadir`. Default is the home directory.

```
[Configuration]
RUNTIME_PATH = GangaNG:GangaAtlas:GangaPanda
gangadir = /scratch/katarzp/gangadir
```

<sup>5</sup>The problem was particularly noticeably during the CSC analysis when the system was still under development.

Users must enable the NorduGrid backend:

```
[ARC]
ARC_ENABLE = True
```

The user tag included in the output location is set by the “usertag”.

```
[DQ2]
usertag = user10
```

Users also need to specify which Virtual Organization they belong to and the Virtual Organization Management System - voms:

```
[LCG]
VirtualOrganisation = atlas

[defaults_GridProxy]
voms = atlas
```

### 5.9.4 NG backend object

Below follows a printout of the object for a successfully finished job:

```
1 backend = NG (
2   status = 'FINISHED' ,
3   check_availability = False ,
4   actualCE = 'siri.lunarc.lu.se' ,
5   submit_options = [] ,
6   middleware = 'ARC' ,
7   CE = 'siri.lunarc.lu.se,pikolit.ijs.si' ,
8   id = 'gsift://siri.lunarc.lu.se:2811/jobs/18601264260301939012271' ,
9   queue = 'arc' ,
10  requestedcputime = '11 hours, 40 minutes' ,
11  reason = None ,
12  cputime = '23 minutes' ,
13  clean = [] ,
14  monInfo = {'remotefile': 'stdout'} ,
15  enable_resume = True ,
16  RejectCE = '' ,
17  requirements = NGRequirements (
18    runtimeenvironment = ['APPS/HEP/ATLAS-15.6.1'] ,
19    move_links_locally = 0 ,
20    gsidcap = 'srm.swegrid.se' ,
21    cputime = 700 ,
22    other = [] ,
23    timeout = 5 ,
24    memory = 3500 ,
25    disk = 3000 ,
26    walltime = 760
```

The most relevant lines are:

- 2 : **status** – ARC status of the job, same as the one shown in the NorduGrid monitor
- 3 : **check\_availability** – test which removes files from the input list if they are not present in the NDGF system. Here it is set to False.
- 4 : **actualCE** – indicates the site where the job is executed
- 7 : **CE** – The user can specify one or more sites she would like as input to the brokering. CE is a comma separated string. One of the required sites will be chosen as **actualCE**. It is not required to use the option. The brokering algorithm will search for the most

suitable resource. However, if the option is used, it is a good practice to specify more than one site.

- 8 : **id** – Unique Grid identification
- 9 : **queue** – Queue to which the job has been submitted
- 10 : **requestedcputime** – CPU time requested in the backend requirements
- 11 : **reason** – For failed jobs, this line gives the Grid error message which is usually a good indication of the reason why the job failed.
- 12 : **cputime** – Shows the CPU time actually spent by the job.
- 15 : **enable\_resume** – In order to have the possibility to rescue jobs using the resume functionality, the enable variable must be set to True. Otherwise the job will be removed from the site after downloading the log files.
- 16 : **RejectCE** – If one wants to take advantage of a broad range of resources and not limit the brokering algorithm by selecting some sites (using **CE**), one can keep the **CE**-field blank and rather reject bad sites instead of selecting good ones. The argument is a comma separated string.
- 17 : **requirements** – This schema allows the user to set some important resource requirements like for example CPU-time or memory consumption. As the ARC middleware respects strictly these and kills the job if they are exceeded, it is useful to have a rough idea of what a job needs.
- 18 : **runtimeenvironment** – ARC sites announce the software installed at a site as so-called run-time environment and specify them by a path, ex. **APPS/MATH/MATLAB-7.5**. A list of run-time environments can be found by clicking on a site in the NorduGrid monitor. In case of ATLAS jobs, this line will be filled automatically during submission. The required ATLAS release corresponds to the one used on the local machine.
- 21 : **cputime** – Computing time required for a job (in minutes).
- 23 : **timeout** – The client brokering surveys the whole NorduGrid collecting updated information from the sites. In case of sites with problems it will wait until a certain timeout. In order to speed up the process one can set a shorter limit. It will eliminate the bad sites while the risk of losing good ones is relatively small. The value is set in seconds.
- 24 : **memory** – Required virtual memory on the processing node. Jobs that exceed the limit are killed by the Grid Manager. The majority of the nodes can currently offer more than 2GB of memory. Setting a high requirement will exclude weaker resources.
- 25 : **disc** – Disc space requested by the job.
- 26 : **walltime** – The wall time a job needs in the system in order to finish.

### 5.9.5 Conclusion

The implementation of the GangaNG backend opened NorduGrid and NDGF resources for analysis to ATLAS physicists. It has shown good performance and high success rate due to stability of the underlying Grid system and the efficient data handling in the NDGF Tier-1.

ATLAS supports also PanDA [100] as an official distributed analysis framework. In order to make resources available to a broader range of users, it is now also possible to submit user analysis jobs to NorduGrid via the PanDA interface. This became possible once the managed production was integrated in the pilot based global Production System. The user jobs are submitted via the same framework. This development brought ATLAS close to the goal of offering two distributed analysis frameworks with equal and full access to all data and computing resources. Naturally, the usage will be regulated by collaboration policies. However, it should not be limited by technical issues.

# Part III

## The Analysis Challenge





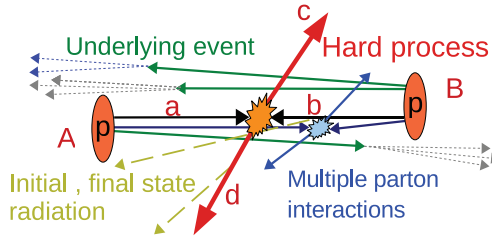
# Chapter 6

## Phenomenology - SUSY and the Standard Model

Chapter 3 gave a general introduction of Supersymmetric extensions of the Standard Model (SM). It showed how the theory is built up and how SUSY models are constructed. The present chapter will first take a closer look at the SUSY phenomenology (Section 6.1) and the constrained mSUGRA model (Section 6.2). Section 6.3 will introduce the various benchmark scenarios studied in the analysis as well as the relevant SM backgrounds. The last two sections give a short description of the MC simulation tools used in preparation of both 14 and 10 TeV samples (Section 6.4) and data preparation procedure (Section 6.5). The following chapters will describe the analysis physics objects as well as SUSY search results at 14 and 10 TeV.

### 6.1 Sparticle production

In a hadron collider the accelerated particles are not elementary and their composite nature makes the processes more complex. The interesting processes occur between partons, gluon and quark constituents, both valence quarks and sea-quarks. Partons carry only a certain fraction  $x$  of the total longitudinal momentum of the colliding hadron.



$$A + B \rightarrow c + d + X$$

Figure 6.1: Hadron collider event.

We start with two hadrons  $A$  and  $B$  involved in the reaction shown in Figure 6.1. The so-called hard process where, for example, SUSY particles  $c$  and  $d$  are produced, is the sub-process  $a + b \rightarrow c + d$  where  $a$  and  $b$  are constituents of the incoming hadrons. The

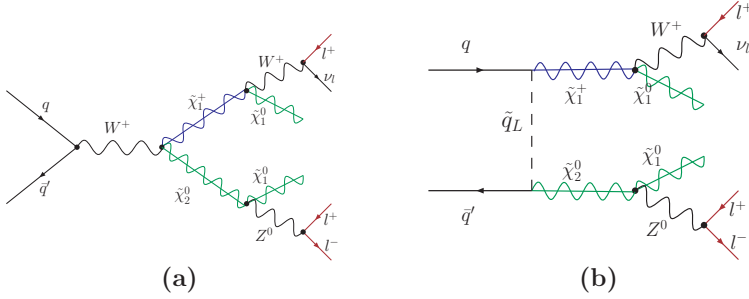


Figure 6.2: Feynman diagrams for direct  $\tilde{\chi}_2^0 \tilde{\chi}_1^\pm$  production with subsequent decay to three leptons, s-channel (a) and t-channel (b).

total picture of the complex interaction involves also initial and final state radiation as well as the beam remnants and multiple parton interactions. The processes accompanying the hard process are also referred to as the underlying event.

The cross section for the hard process is obtained by convoluting the production cross section for the sub-process with the parton distribution functions of the involved constituent particles  $a$  and  $b$  [60]. The following section will focus on some of the main sparticle production processes at LHC.

### 6.1.1 Gaugino production

At the Tevatron which is a  $p\bar{p}$  collider with a center of mass energy of 2 TeV, chargino and neutralino production is expected to have the largest SUSY cross section. These are produced through quark-antiquark annihilation. As Tevatron collides  $p\bar{p}$  these are found as valence quarks, while at LHC such pairs must involve sea-quarks. The energy reach at the Tevatron may be too low for production of squarks and gluinos, favouring the electroweak processes, while at LHC, as soon as the QCD processes are kinematically open, they will dominate. Nevertheless, gaugino production is expected at LHC in many SUSY scenarios, particularly in those where the supersymmetric fermions and possibly also gluinos are very heavy. In cases where the strongly interacting SUSY particles are beyond the LHC energy reach such electroweak processes might even become discovery channels.

In most common SUSY scenarios where the gaugino mass unification implies approximately  $M_1 : M_2 : M_3 \approx 1 : 2 : 7$  and  $\mu \gg M_1, M_2$ , the neutralino  $\tilde{\chi}_1^0$  is mostly bino-like, while the heavier  $\tilde{\chi}_2^0$  and chargino  $\tilde{\chi}_1^\pm$  are mostly wino-like and have therefore large couplings to  $W$  and the squark-quark system. The heaviest gauginos  $\tilde{\chi}_3^0$ ,  $\tilde{\chi}_4^0$  and  $\tilde{\chi}_2^\pm$  are then mainly higgsino like. The composition of the mass eigenstates is crucial for the production and decay of these particles as it defines their couplings.

The gaugino production with highest cross section is often the pair production of a  $\tilde{\chi}_1^\pm \tilde{\chi}_2^0$  pair which can occur in a quark-antiquark annihilation, via  $W$  exchange in the s-channel and left-handed squark exchange in the t-channel, see Figure 6.2.

The chargino pair production occurs via quark-antiquark annihilation. The production of the lightest pair  $\tilde{\chi}_1^\pm \tilde{\chi}_1^\mp$  has the highest cross section and together with  $\tilde{\chi}_2^\pm \tilde{\chi}_2^\mp$  they are produced via  $\gamma$  or  $Z$  exchange in the s-channel or via squark exchange in the t-channel. Figure 6.3 (a) and (b) show the Feynman diagrams for these processes. The production of

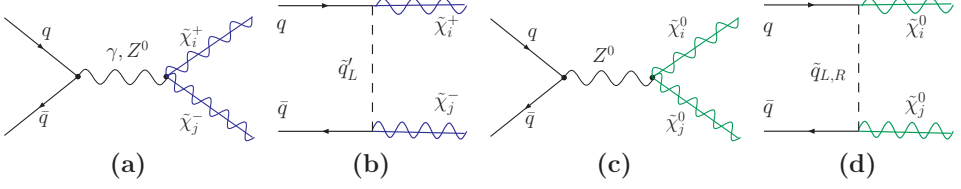


Figure 6.3: Feynman diagrams for direct of  $\chi_i^\pm \chi_j^\mp$  and  $\chi_i^0 \chi_j^0$  pairs. Graphs (a) and (c) show the  $s$ -channel, graphs (b) and (c)  $t$ -channel.

mixed pairs like  $\tilde{\chi}_1^\pm \tilde{\chi}_2^\mp$  can only be produced via  $Z$  exchange in the  $s$ -channel, in addition to the left-handed squark exchange  $t$ -channel.

The neutralino pairs may also be produced in quark-antiquark annihilation with  $Z$  exchange in the  $s$ -channel, see Figure 6.3 (c). The squarks that are exchanged in the  $t$ -channel can now be both left- and right-handed, Figure 6.3 (d). In  $pp$ -collisions  $\tilde{\chi}_1^0$  and  $\tilde{\chi}_2^0$  are mostly strongly produced.

Another electroweak SUSY production process is the creation of slepton and sneutrino pairs. However, as it has very low cross sections and is not relevant in this analysis, it will be omitted.

### 6.1.2 Squark and gluino production

In hadron colliders, QCD processes will dominate if they are not kinematically suppressed. Gluino pairs can be produced via gluon-gluon and quark-antiquark scattering. In Figure 6.4 the main leading order diagrams are shown. As the gluon luminosity falls rapidly off with growing  $x$ , the gluino production through gluon scattering ( $gg \rightarrow \tilde{g}\tilde{g}$  see Figure 6.4 (a)) will be dominant for small  $m_{\tilde{g}}$ , while the quark scattering ( $q\bar{q} \rightarrow \tilde{g}\tilde{g}$ , see Figure 6.4 (b)) and gluon-quark ( $gq \rightarrow \tilde{g}\tilde{q}$ , see Figure 6.7) will dominate for higher gluino masses.

Squark pairs  $\tilde{q}\tilde{q}$  can be produced in gluon scattering processes via diagrams like those in Figure 6.5. Such processes lead to pairs of same flavor and chirality. If kinematically allowed, gluon scattering can lead to pairs of all flavors. While quark scattering as shown in Figure 6.6 contributes mainly to production of flavors with significant luminosity in

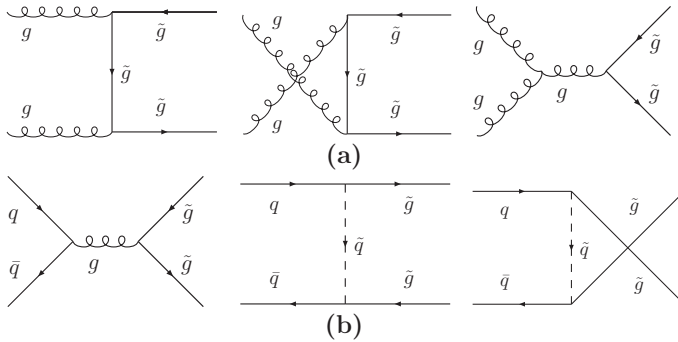


Figure 6.4: Feynman diagrams for gluino production: gluon-scattering (a) and quark-scattering (b).

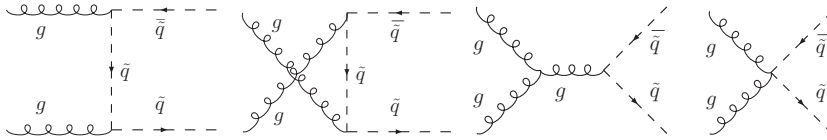


Figure 6.5: Feynman diagrams for squark pair production via gluon scattering.

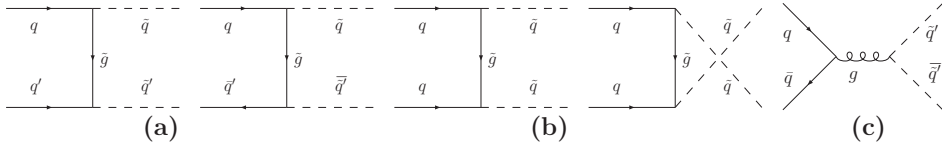


Figure 6.6: Feynman diagrams for squark pair production via quark scattering.

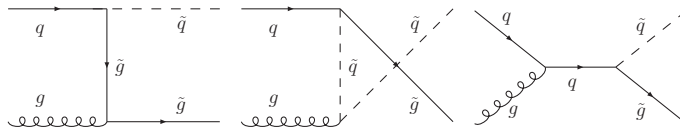


Figure 6.7: Feynman diagrams for gluino-squark production.

the colliding hadrons.

Figure 6.7 shows the Feynman diagrams for the gluon-squark production via gluon-quark scattering. In scenarios where  $m_{\tilde{g}} \approx m_{\tilde{q}}$ , the  $\tilde{g}\tilde{q}$  production dominates over a wide range of  $m_{\tilde{g}}$ . This is the case for three of the ATLAS benchmark points which will be introduced in the following sections. The squark-gluino production accounts there for the largest fraction of the total cross section as shown in Tables 9.1 and 9.2.

Only for very heavy gluinos will the  $\tilde{q}\tilde{q}$  production have a higher cross section. In scenarios where the squarks are significantly heavier than gluinos, the production of these lightest strongly interacting particles will dominate at all mass scales. Figure 6.8 shows the cross section for squark and gluino production [60].

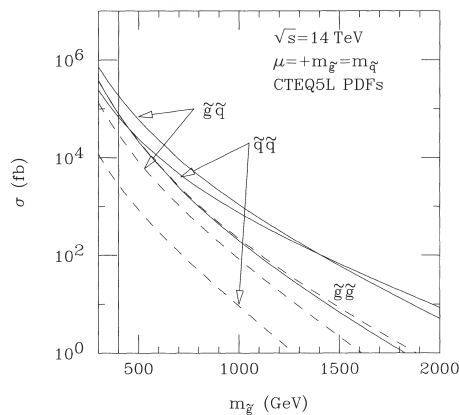


Figure 6.8:

Squark and gluino production cross sections as function of  $m_{\tilde{g}}$  in 14 TeV  $pp$ -collisions for  $m_{\tilde{q}} \approx m_{\tilde{g}}$  (solid line) and  $m_{\tilde{q}} \approx 2m_{\tilde{g}}$  (dashed line) [60].

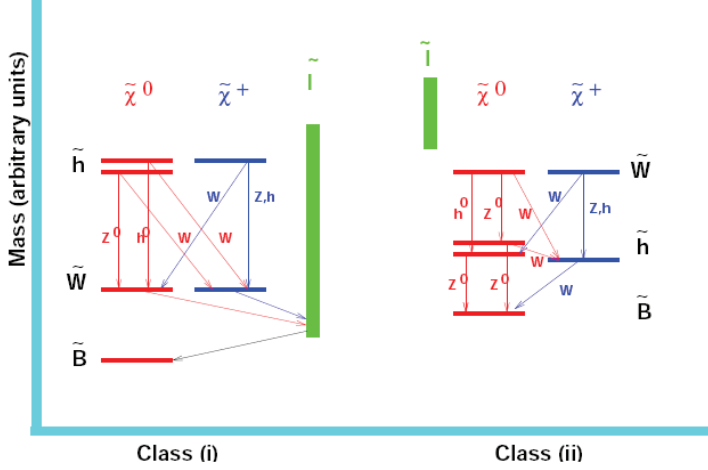


Figure 6.9: Typical decay patterns for two common benchmark scenarios. [138]

### 6.1.3 Gaugino decays

The gaugino decay patterns are defined by their gauge eigenstate mixture and the kinematical constraints due to mass differences. Figure 6.9 illustrates the most commonly studied benchmark points which can be divided into two classes [138]:

- Case (i),  $\mu \gg M_1, M_2$ 
  - The lightest neutralino  $\tilde{\chi}_1^0$  is bino-like ( $\tilde{B}^0$ )
  - Neutralino  $\tilde{\chi}_2^0$  is wino-like ( $\tilde{W}^0$ )
  - The two heaviest neutralinos  $\tilde{\chi}_3^0$  and  $\tilde{\chi}_4^0$  are higgsino like
  - Chargino  $\tilde{\chi}_1^\pm$  is wino-like, while  $\tilde{\chi}_2^\pm$  is higgsino like
  - In addition, the constraint  $m_0 \leq m_{1/2}$  (mSUGRA parameters) makes at least one of the sleptons lighter than  $\tilde{\chi}_2^0$  and  $\tilde{\chi}_1^\pm$
- Case (ii) which corresponds to so called Focus-point scenarios,  $M_1 < |\mu| < M_2$ :
  - The lightest neutralino  $\tilde{\chi}_1^0$  is still mostly bino.
  - Neutralinos  $\tilde{\chi}_2^0$  and  $\tilde{\chi}_3^0$  are more higgsino like
  - Neutralino  $\tilde{\chi}_4^0$  is wino-like
  - Chargino  $\tilde{\chi}_1^\pm$  is now more higgsino-like while  $\tilde{\chi}_2^\pm$  have a larger wino admixture.
  - This parameter space region is characterized by very heavy squarks and sleptons. In some models also the gluino may be heavy.

With this schema as general introduction, we can now take a closer look at the gaugino decays. These can in general be divided into two-body decays which dominate if kinematically accessible, and three-body decays via off-shell particles.

**Charginos** can only decay via the electroweak interaction and the two-body decays like those listed below dominate whenever kinematically allowed [60]:

$$\tilde{\chi}_j^+ \rightarrow W^+ \tilde{\chi}_i^0, H^+ \tilde{\chi}_i^0 \quad (6.1)$$

$$\rightarrow \tilde{\nu}_e \bar{e}, \tilde{e}_L \nu_e, \text{ and the other generations} \quad (6.2)$$

$$\rightarrow \tilde{u}_L \bar{d}, \tilde{d}_L u, \text{ and the other generations} \quad (6.3)$$

$$\tilde{\chi}_2^+ \rightarrow Z \tilde{\chi}_1^+, h \tilde{\chi}_1^+, H \tilde{\chi}_1^+, A \tilde{\chi}_1^+. \quad (6.4)$$

These decay channels are very typical in scenarios corresponding to Case (i) in Figure 6.9. In most common cases the squarks are heavier than the gauginos so the decay to slepton is most relevant. Considering the decay channels which involve gauginos, for the light chargino the decay to  $\tilde{\chi}_1^0$  (Eq. 6.1) is most typical, while the heavier chargino may also decay to  $\tilde{\chi}_1^+$  like in Eq. 6.4 starting a small cascade.

If the two-body decays are kinematically inaccessible ( $m_{\tilde{\chi}_1^+} < m_W + m_{\tilde{\chi}_1^0}$ ), three-body decays

$$\tilde{\chi}_j^+ \rightarrow \tilde{\chi}_i^0 + f \bar{f}' \quad (6.5)$$

which in most cases occur via a virtual boson will dominate. Feynman diagram for such decay can be found in Figure 6.2. The branching fractions will then be very similar to  $B(W \rightarrow f \bar{f}')$ . Such decays are common in scenarios corresponding to Case(ii) in Figure 6.9 where the sleptons are very heavy.

In three-body decays like in Eq. 6.5 the contribution from the first two generation slepton or sneutrino exchange will usually be small. However, for large  $\tan \beta$  the  $\tilde{\tau}_1$  becomes light and starts to play an important role.  $\tilde{\tau}_1$  can be involved in the three-body decay, while for very high  $\tan \beta$  it may even become accessible via two-body decay  $\tilde{\chi}_j^+ \rightarrow \tilde{\tau}_1 \nu_\tau$  and become then the dominating channel.

**Neutralinos** will, if kinematically allowed, undergo two-body decays [60]:

$$\tilde{\chi}_i^0 \rightarrow W^+ \tilde{\chi}_j^-, (H^+ \tilde{\chi}_j^-), Z \tilde{\chi}_{i'}^0, h \tilde{\chi}_{i'}^0, (A \tilde{\chi}_{i'}^0, H^0 \tilde{\chi}_{i'}^0) \quad (6.6)$$

$$\rightarrow \tilde{l}_{L,R} \bar{l}, \tilde{\nu}_l \bar{\nu}_l, (\tilde{q}_{L,R} \bar{q}), \text{ (and charge conserving combinations)} \quad (6.7)$$

Decays in brackets are in most scenarios kinematically suppressed. Examples of such scenarios are illustrated in Cases (i) in Figure 6.9. If the two-body decays are closed, three-body decays like

$$\tilde{\chi}_i^0 \rightarrow \tilde{\chi}_{i'}^0 + f \bar{f}, \quad \tilde{\chi}_i^0 \rightarrow \tilde{\chi}_j^+ + f \bar{f}' \quad (6.8)$$

via off-shell gauge bosons will be most relevant (Case (ii)). Feynman diagram of three-body decay can be found in Figure 6.2.

Taking  $\tilde{\chi}_2^0$  as example [60], we observe that for light  $\tilde{\chi}_2^0$  final states with  $q \bar{q} \tilde{\chi}_1^0$  will dominate, while the two-body decay to  $Z^0 \tilde{\chi}_1^0$  dominates as soon as it becomes kinematically allowed. For even higher masses of  $\tilde{\chi}_2^0$  the decay channel involving the light Higgs  $h^0 \tilde{\chi}_1^0$  will be the most important.

Like in the case of charginos, the effects of  $\tan \beta$  on the mass hierarchy of the third generation sfermions is mirrored in the neutralino decay patterns. For low and intermediate values of  $\tan \beta$  the final states with quarks  $u, d, s, c$  and  $b$  dominate. The first generation leptons account for only a very small fraction, while the  $\tilde{\tau}_1$  become increasingly important for higher  $\tan \beta$ . First it enters in three-body decays and when the  $\tilde{\tau}_1$  becomes very light, the two-body decays to  $\tilde{\tau}_1 \tau$  dominate.

### 6.1.4 Gluino and squark decays

The trilepton signal which is the focus of this analysis may also occur in association with jets in cascade decays initiated by gluinos and squarks. Gluinos can only decay via strong interaction where the  $q\tilde{q}_{L,R}$  channel dominates if kinematically allowed and the gluino-squark-quark coupling has QCD strength  $\propto \sqrt{\alpha_S}$ . As the third generation squarks may be the lightest ones, they may be the only accessible two-body decay mode. If all squarks are heavier than the gluino, the latter will decay via off-shell squarks into  $q\tilde{q}\tilde{\chi}_i^0$  or  $q\tilde{q}'\tilde{\chi}_i^\pm$  states.

In many scenarios we observe that the kinematically favoured decay  $q\tilde{q}\tilde{\chi}_1^0$  has lower branching fraction than  $q\tilde{q}\tilde{\chi}_2^0$  and  $q\tilde{q}'\tilde{\chi}_1^\pm$ . This can be explained by the composition of the gauginos which in the common scenarios where  $2M_1 \simeq M_2 \simeq m_{\tilde{g}}/3 \ll \mu$  follow Case (i) in Figure 6.9. The heavier gauginos are favoured as the  $SU(2)_L$  coupling  $g = e/\sin\theta_W$  is stronger than the  $U(1)_Y$  hypercharge coupling  $g' = e/\cos\theta_W$ , resulting in a suppression of the decay to the bino-like  $\tilde{\chi}_1^0$  by a factor of  $\tan\theta_W \approx 0.54$ . The fact that the mass eigenstates are mixtures of gauge eigenstates makes of course the picture more complex. In scenarios with heavier  $\tilde{g}$  the gaugino compositions tend to follow Figure 6.9, Case(ii), and decays involving the wino-like  $\tilde{\chi}_2^\pm$  and  $\tilde{\chi}_4^0$  become important.

The  $\tilde{g}$  branching fractions vary also with  $\tan\beta$  and as it becomes larger, decays to third generation squarks and quarks, especially b-quarks, become important.

**Squarks** will also preferably decay via two-body decays [60]

$$\tilde{u}_L \rightarrow u\tilde{\chi}_i^0, d\tilde{\chi}_i^+, u\tilde{g} \quad \tilde{d}_L \rightarrow d\tilde{\chi}_i^0, u\tilde{\chi}_i^-, d\tilde{g} \quad (6.9)$$

$$\tilde{u}_R \rightarrow u\tilde{\chi}_i^0, u\tilde{g} \quad \tilde{d}_R \rightarrow d\tilde{\chi}_i^0, d\tilde{g} \quad (6.10)$$

If kinematically allowed, the  $q\tilde{g}$  QCD decay will dominate, otherwise the electroweak decay involving gauginos takes over. Only the left-handed squarks couple to  $\tilde{\chi}_i^\pm$ , which in most cases, is wino like. By the same argument ( $g > g'$ ), left-handed squarks preferably decay to the heavier  $\tilde{\chi}_1^\pm$  and  $\tilde{\chi}_2^0$  rather than to the bino-like  $\tilde{\chi}_1^0$ . The decays to higgsino-like gauginos ( $\tilde{\chi}_3^0, \tilde{\chi}_4^0, \tilde{\chi}_2^\pm$ , Case (i) in Figure 6.9) are only important for stops and sbottoms which have a larger Yukawa coupling [40]. For higher squark masses when the schema is changed towards Case (ii), decays to the wino-like  $\tilde{\chi}_4^0$  and  $\tilde{\chi}_2^\pm$  dominate. For the right-handed squarks only the decay involving  $\tilde{\chi}_1^0$  and  $\tilde{\chi}_2^0$  are relevant until the  $q\tilde{g}$  channel is open.

The stop which may be relatively light will preferably decay to  $t\tilde{g}$  or  $b\tilde{\chi}_1^\pm$ . However, if these are kinematically forbidden stops may be forced to decay via flavor violation decays like  $\tilde{t}_1 \rightarrow c\tilde{\chi}_1^0$ .

Third generation squarks are special due to strong mixing-effects and non-negligible Yukawa couplings. Their decay patterns are more complicated and include also the Higgs sector.

### 6.1.5 Slepton decays

Finally, some words about slepton decays. The direct slepton production has a relatively low cross section and in this analysis they are most interesting as gaugino decay products. Slepton decays follow a similar pattern as the electroweak squark decays [60]

$$\tilde{e}_L \rightarrow e\tilde{\chi}_i^0, \nu_e\tilde{\chi}_i^- \quad \tilde{\nu}_{eL} \rightarrow \nu_e\tilde{\chi}_i^0, e\tilde{\chi}_i^+ \quad (6.11)$$

$$\tilde{e}_R \rightarrow e\tilde{\chi}_i^0 \quad (6.12)$$

Only the left-handed sleptons couple to the wino-like gauginos ( $\tilde{\chi}_2^0$  and  $\tilde{\chi}_1^+$ ) which then dominate the cross section if they are kinematically allowed. For very heavy slepton scenarios where the gaugino mixing schema moves towards Case (ii) in Figure 6.9, decays involving  $\tilde{\chi}_4^0$  and  $\tilde{\chi}_2^+$  dominate.

Like in the case of third generation squarks, the stau family decays are more complicated and if kinematically allowed, the Higgs sector is also involved.

## 6.2 Constrained MSSM – mSUGRA

The present work is part of the recent ATLAS analysis effort, the so called Computer System Commissioning (CSC) which is summarized in the ATLAS CSC-Book [78] (p. 1603–1605 and p. 1643–1659). In this analysis ATLAS has focused on developing strategies based on the mSUGRA model.

The five mSUGRA parameters  $m_{1/2}$ ,  $m_0$ ,  $A_0$ ,  $\tan\beta$  and  $\text{sign}(\mu)$  span a multidimensional plane of possible scenarios. However, a number of theoretical and experimental results, as well as astrophysical observations severely reduce the model to a small set of characteristic phenomenological scenarios, which are mainly motivated by consistency with the dark matter relic density measurements [49, 50, 52].

The LSPs would in the early universe be in thermal equilibrium, meaning that they would be created and annihilated at equal rate. As the universe expanded and cooled down the production stopped and the annihilation became dominant leading to a steady decrease of the LSP density. Then at some point in the expansion of the universe the density became too small and the collision rate became negligible. This process is called the LSP freeze-out. The remaining density of LSPs is determined by the cross section of the annihilation processes ( $\tilde{\chi}_1^0\tilde{\chi}_1^0 \rightarrow f\bar{f}, W^+W^-, ZZ$ ), in some models also co-annihilation ( $\tilde{\chi}_1^0\tilde{\tau} \rightarrow \tau\gamma$ ), before the freeze-out. The phenomenological benchmark points are therefore characterized by various features that lead to dark matter density consistent with measurements.

The following are the distinctions of the four studied points<sup>1</sup>. Further details are given in the following sections.

- SU1: Coannihilation region – where  $m_{\tilde{\tau}_1} \sim m_{\tilde{\chi}_1^0}$  such that  $\tilde{\chi}_1^0\tilde{\tau} \rightarrow \tau\gamma$  co-annihilation in the early universe reduces the relic density
- SU2: Focus-point region –  $m_{1/2} \ll m_0$  - the lightest neutralino has a significant Higgsino component which enhances the  $\tilde{\chi}_1^0\tilde{\chi}_1^0 \rightarrow W^+W^-$  annihilation cross section
- SU3: Bulk region – LSP annihilation happens through the exchange of light sleptons or squarks ( $\tilde{\chi}_1^0\tilde{\chi}_1^0 \rightarrow f\bar{f}$ )
- SU4: Low mass point – with a low value of the SUSY mass scale, close to Tevatron models ( $\tilde{\chi}_1^0\tilde{\chi}_1^0 \rightarrow f\bar{f}$ )

For a quick reference, Table 6.1 gives an overview of the points studied in this analysis while Figure 6.10 gives a graphical overview of the parameter space.

---

<sup>1</sup>ATLAS operates with more than these four points, however, these are the most relevant for trilepton analysis.



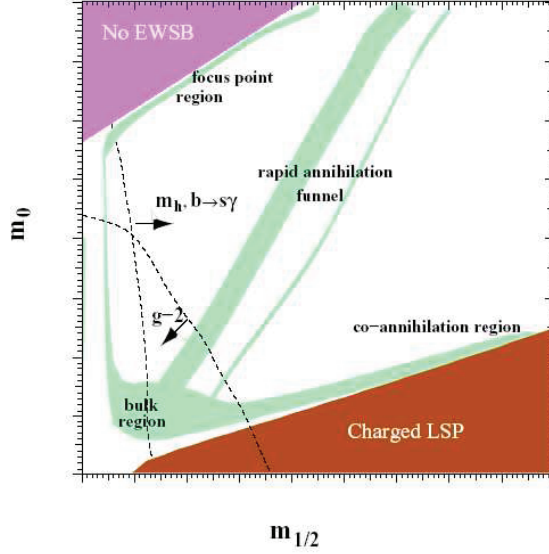


Figure 6.10: The mSUGRA parameter space spanned by  $m_{1/2}$  and  $m_0$ . The green/light gray regions are consistent with the observed dark matter relic density. The  $\Omega_\chi h^2$  interval used in the plot is much broader than the current limits and gives only a conceptual picture. (see Section 3.9)

| Point | $m_0$ | $m_{1/2}$ | $A_0$ | $\tan \beta$ | $\text{sign} \mu$ | Point Characteristics |
|-------|-------|-----------|-------|--------------|-------------------|-----------------------|
| SU1   | 70    | 350       | 0     | 10           | +                 | Coannihilation region |
| SU2   | 3550  | 300       | 0     | 10           | +                 | Focus point region    |
| SU3   | 100   | 300       | -300  | 6            | +                 | Bulk region           |
| SU4   | 200   | 160       | -400  | 10           | +                 | Low mass point        |

Table 6.1: Summary of parameters defining the ATLAS mSUGRA benchmark points. The list contains only points studied in this analysis.

## 6.3 Signal and background in mSUGRA searches

This section gives a more detailed presentation of the particle spectra and phenomenological characteristics of the studied mSUGRA benchmark scenarios. The first part of the analysis is done with fully simulated events at 14 TeV (official CSC-samples). These MC samples were defined long before the 19th September 2008 incident and the revised LHC operation plans with lower center of mass energies.

For the CSC SUSY MC-samples the particle masses and branching fractions are calculated using the ISASUSY [139] module from ISAJET [140] with a top mass of 175 GeV. The rest of the event is simulated using HERWIG [141] in combination with JIMMY [142]. The K-factors are calculated using PROSPHINO [143] and the parton distribution set CTEQ6M [144].

| Particle           | SU1    | SU2     | SU3    | SU4    |
|--------------------|--------|---------|--------|--------|
| $\tilde{d}_L$      | 764.90 | 3564.13 | 636.27 | 419.84 |
| $\tilde{u}_L$      | 760.42 | 3563.24 | 631.51 | 412.25 |
| $\tilde{b}_1$      | 697.90 | 2924.80 | 575.23 | 358.49 |
| $\tilde{t}_1$      | 572.96 | 2131.11 | 424.12 | 206.04 |
| $\tilde{d}_R$      | 733.53 | 3576.13 | 610.69 | 406.22 |
| $\tilde{u}_R$      | 735.41 | 3574.18 | 611.81 | 404.92 |
| $\tilde{b}_2$      | 722.87 | 3500.55 | 610.73 | 399.18 |
| $\tilde{t}_2$      | 749.46 | 2935.36 | 650.50 | 445.00 |
| $\tilde{e}_L$      | 255.13 | 3547.50 | 230.45 | 231.94 |
| $\tilde{\nu}_e$    | 238.31 | 3546.32 | 216.96 | 217.92 |
| $\tilde{\tau}_1$   | 146.50 | 3519.62 | 149.99 | 200.50 |
| $\tilde{\nu}_\tau$ | 237.56 | 3532.27 | 216.29 | 215.53 |
| $\tilde{e}_R$      | 154.06 | 3547.46 | 155.45 | 212.88 |
| $\tilde{\tau}_2$   | 256.98 | 3533.69 | 232.17 | 236.04 |
| $\tilde{g}$        | 832.33 | 856.59  | 717.46 | 413.37 |
| $\tilde{\chi}_1^0$ | 136.98 | 103.35  | 117.91 | 59.84  |
| $\tilde{\chi}_2^0$ | 263.64 | 160.37  | 218.60 | 113.48 |
| $\tilde{\chi}_3^0$ | 466.44 | 179.76  | 463.99 | 308.94 |
| $\tilde{\chi}_4^0$ | 483.30 | 294.90  | 480.59 | 327.76 |
| $\tilde{\chi}_1^+$ | 262.06 | 149.42  | 218.33 | 113.22 |
| $\tilde{\chi}_2^+$ | 483.62 | 286.81  | 480.16 | 326.59 |
| $h^0$              | 115.81 | 119.01  | 114.83 | 113.98 |
| $H^0$              | 515.99 | 3529.74 | 512.86 | 370.47 |
| $A^0$              | 512.39 | 3506.62 | 511.53 | 368.18 |
| $H^+$              | 521.90 | 3530.61 | 518.15 | 378.90 |
| $t$                | 175.00 | 175.00  | 175.00 | 175.00 |

Table 6.2: Particle mass spectrum (in GeV) for the SUSY benchmark points. Adapted from [78]

### 6.3.1 Co-annihilation Region (SU1)

SU1 is a benchmark point in the co-annihilation region which is characterized by low  $\tilde{\tau}_1$  mass. The gluino and squarks have a slightly higher mass than in the low mass points, however at 14 TeV it has a considerable cross section of 8.2 pb (LO). The main SUSY process is production of  $\tilde{q}_{L/R}\tilde{g}$  pairs. A detailed listing of the sparticle decays can be found in Table B.1 (Appendix B). As expected, the light stau is involved in many decays.

### 6.3.2 The Focus Point Region (SU2)

This region of the parameter space is characterized by a very high value of  $m_0$  which pushes the masses of the sfermions above  $\sim 3$  TeV. The common sfermion mass  $m_{1/2}$  is on the other hand relatively low, thus making the gauginos and the gluino the lightest sparticles in the spectrum, as can be seen in Figure 6.11 (a). Table 6.2 lists the sparticle masses, and clearly shows the specific mass hierarchy of this parameter space region. The direct production of gauginos accounts for 89.8% of the total cross section which is 5.2 pb

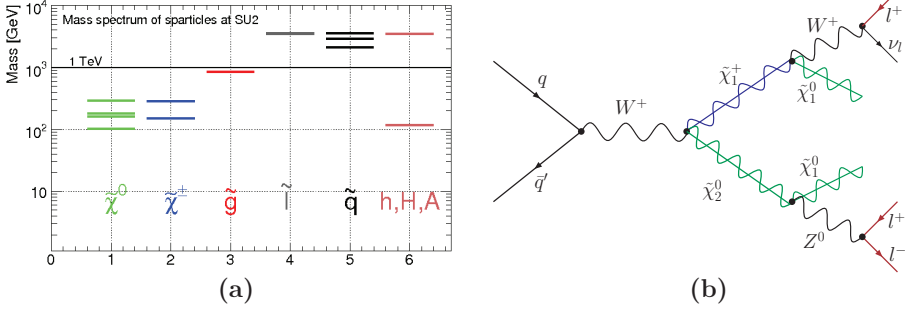


Figure 6.11: (a): SU2 sparticle mass spectrum. (b): Feynman diagram (s-channel) representing the direct production of a chargino and neutralino pair which dominates the SU2 cross section.

(LO). Pairs of chargino-neutralino or neutralino-neutralino that can lead to three or more leptons in the initial state represent 42.3% of the total cross section. A Feynman diagram for the dominant s-channel  $W$  exchange leading to a  $\tilde{\chi}_1^\pm \tilde{\chi}_2^0$  is shown Figure 6.11 (b). Of all processes that can lead to trilepton final state, this has the highest cross section. In addition there is a noticeable gluino pair production cross section (10.1%). Table B.2 (Appendix B) lists the main gaugino and gluino branching fractions. We can see that although the cross section for  $\tilde{\chi}_1^\pm \tilde{\chi}_2^0$  pair production is expected to be high, the gaugino branching fractions to leptons are relatively low. The  $\tilde{\chi}_2^0$  and  $\tilde{\chi}_1^\pm$  decay via virtual  $Z$  and  $W$  and their decay fractions resemble the patterns found for the gauge bosons, namely  $W \rightarrow l\nu \simeq 11\%$ ,  $Z \rightarrow l^+l^- \simeq 3.4\%$  per flavor.

### 6.3.3 The Bulk Point Region (SU3)

This point represents a very typical mSUGRA scenario with relatively light mass spectrum as presented in Table 6.2 and in Figure 6.12 (a). The LO cross section is 20.7 pb and the dominant production processes are  $qg \rightarrow \tilde{q}_{R/L} \tilde{g} \simeq 44\%$ ,  $qq \rightarrow \tilde{q} \tilde{q} \simeq 36\%$  and  $gg \rightarrow \tilde{g} \tilde{g} \simeq 8\%$ . The gluino is the heaviest colored sparticle and can initiate cascade decays via squarks like the one shown in Figure 6.12 (b). A typical SU3 event has many, often more than four, relatively hard jets. In this benchmark scenario which corresponds to Class (i) in Figure 6.9,  $\tilde{\tau}_1$  is the lightest slepton ( $m_{\tilde{\tau}_1} = 150$  GeV), significantly lighter than  $\tilde{\chi}_1^\pm$  and  $\tilde{\chi}_2^0$ , which therefore has a high branching fraction for channels involving the third (s)lepton family. As these lighter gauginos are common in squark decays, a large fraction of the SU3 events involve  $\tau$ 's. Table B.3 (Appendix B) lists the main branching fractions of the particles in SU3.

### 6.3.4 The Low Mass Point Region (SU4)

This point is also in the Bulk region, but it is placed very close to the scenarios that are currently being probed at the Tevatron. It is characterized by a very light particle spectrum, summarised in Figure 6.13, and it has the highest cross section of all ATLAS benchmark points: 294.5 pb (LO). In addition, there is a change in the mass hierarchy compared to SU3. A number of squarks ( $\tilde{t}_1, \tilde{b}_1, \tilde{u}_R, \tilde{d}_R$  and even  $\tilde{u}_L$ ) are lighter than the gluino. The dominant production processes are  $qg \rightarrow \tilde{q}_{R/L} \tilde{g} \simeq 45\%$ ,  $qq \rightarrow \tilde{q} \tilde{q} \simeq 31\%$  and

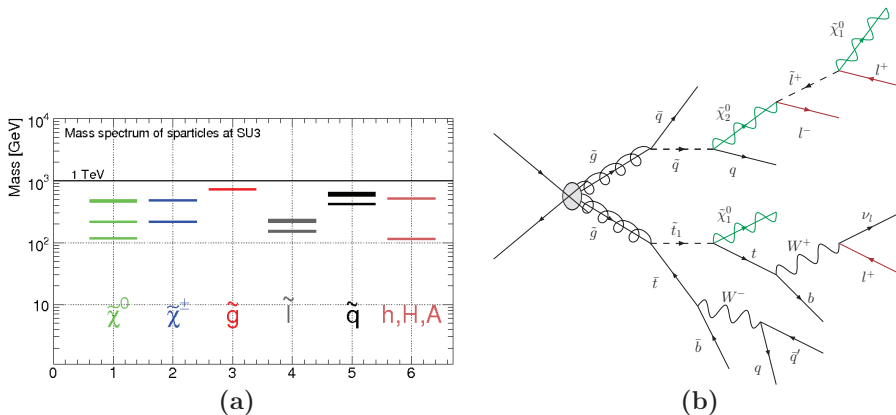


Figure 6.12: (a): SU3 sparticle mass spectrum. (b): Feynman diagram for a gluino production followed by a cascade decay with three leptons in the final stat.

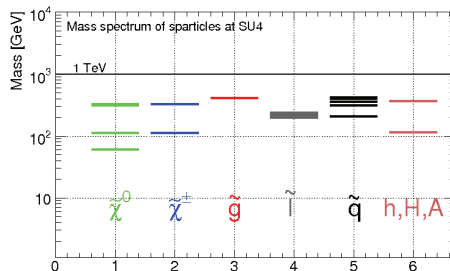


Figure 6.13: SU4 sparticle mass spectrum.

$$gg \rightarrow \tilde{g}\tilde{g} \simeq 17\%.$$

The SU4 gaugino decay pattern differs from SU3 as the tau plays a less important role and there are more final states involving quarks. A detailed listing of branching fractions can be found in Table B.4 (Appendix B). The trilepton final states are mainly due to cascade decays like the one showed in Figure 6.12 (b).

### 6.3.5 Standard Model Background

Since LHC is a  $pp$ -collider, QCD processes will have the highest cross sections. Presence of isolated leptons, on the other hand, is a signature of many interesting and possibly new processes and it is exploited in many search strategies. Requiring as many as three isolated leptons in the final state reduces significantly the background from QCD or other SM processes as multi-lepton final states are rare.

Nevertheless, there is a number of so-called reducible backgrounds, meaning processes with two or more leptons in final state with additional extra leptons originating from jets of photons. With a cross section of 449.8 pb (NLO, only leptonic modes)  $t\bar{t}$  production represents one of the most dominant trilepton backgrounds. The MC sample used in this analysis contains no all-hadronic events so that one or two isolated leptons were expected

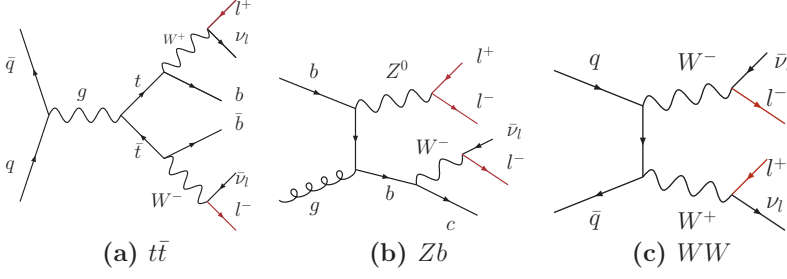


Figure 6.14: SM backgrounds

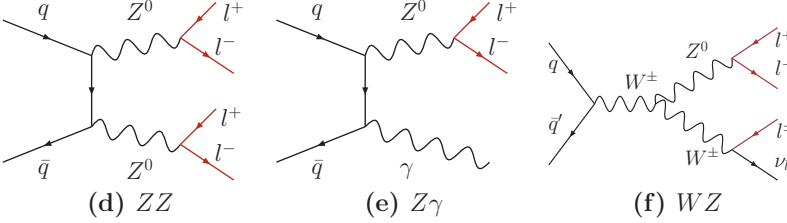


Figure 6.15: Feynman diagrams for some of the main SM backgrounds leading to three leptons in final state. All processes except (f)  $WZ$  are reducible backgrounds which obtain three leptons in final state due to additional effects.

in all events. A Feynman diagram for a leptonic  $t\bar{t}$  event is shown in Figure 6.14 (a). In a  $t\bar{t}$  event there are always two  $b$ -jets which often contain leptons from semileptonic  $b$ -decays. Such leptons are typically surrounded by significant energy deposits in the calorimeters and tracks, however it is not unusual that they pass isolation criteria or other discriminating requirements. In addition there may be fake leptons which in reality are misidentified jets.

The situation is similar in  $Zb$ -jet events, Figure 6.14 (b), where there may also be additional leptons from semileptonic  $b$ -decays. The  $b$ -jet itself can come from the hard process or from initial state radiation [145]. The process has a cross section of 205 pb (NLO) which is considerably higher than most SUSY processes, see Table 6.3.

At the LHC there is a very rich spectrum of processes involving a gauge boson and jets,  $Zb$  is only one of them, however for the present analysis at 14 TeV only this sample could be used as the others were imposed kinematical cuts on jets which were incompatible with the trilepton event selection. As the  $Zb$  events are expected to give a significant contribution to the SM background it was considered a good representative for these processes although it most probably gives an underestimated background. A broader spectrum of background samples will be included in the analysis presented in Chapter 9.

Another class of events which is expected to contribute is the gauge boson production,  $WW$ ,  $ZZ$ ,  $Z\gamma$  and  $WZ$  which is illustrated in Figure 6.14 (c) and Figure 6.15 (d)-(f). Although the cross sections are moderate, the high number of isolated leptons and low hadronic activity make them particularly challenging in searches for SUSY trilepton events in the jet exclusive mode. In  $WW$  one expects no more than two isolated leptons, however there may be contributions from initial state radiation.

The lepton multiplicity cut requires three or more leptons. A  $ZZ$  event can therefore be accepted as a four-lepton event, or as trilepton event due to lepton reconstruction inefficiencies. In  $Z\gamma$  events there may be additional two electrons from  $\gamma$ -conversion and as with  $ZZ$  background they can be accepted as three- or four-lepton events.

The  $WZ$  production is a SM counterpart of the process in Figure 6.11 (b) and may be considered an irreducible background. All diboson samples, except  $Z\gamma$ , were simulated including effects of the off-shell bosons. The  $Z\gamma$  sample was imposed a  $p_T > 25$  GeV cut on the photon (and  $|\eta| < 2.7$ ) which could possibly have given an underestimated background.

The cross section for both the signal and background can be found in Table 6.3. Column three ( $\sigma_{LO} \times \epsilon_F[pb]$ ) gives the LO cross sections multiplied with the generator filter efficiency. Column four gives the average event weight,  $\langle w \rangle$ , which is different from one for samples including higher order corrections. Column five ( $\sigma_{NLO} \times \epsilon_F$ ) shows the NLO cross sections which in most cases are obtained through  $K$ -factors. Finally the right most column,  $\int dt \mathcal{L}$ , shows the generated integrated luminosity of the MC samples.

| Process    | sample # | $\sigma_{LO} \times \epsilon_F [pb]$ | $\langle w \rangle$ | $\sigma_{NLO} \times \epsilon_F [pb]$ | $\int dt \mathcal{L} [fb^{-1}]$ |
|------------|----------|--------------------------------------|---------------------|---------------------------------------|---------------------------------|
| SU1        | 005401   | 8.2                                  | 1                   | 10.9                                  | 18.3                            |
| SU2        | 005402   | 5.2                                  | 1                   | 7.2                                   | 6.6                             |
| SU3        | 005403   | 20.9                                 | 1                   | 27.7                                  | 17.0                            |
| SU4        | 005404   | 294.5                                | 1                   | 403.8                                 | 0.48                            |
| $WW$       | 005985   | 24.5                                 | 1                   | 40.9                                  | 1.2                             |
| $WZ$       | 005987   | 7.8                                  | 1                   | 15.99                                 | 3.0                             |
| $ZZ$       | 005986   | 2.1                                  | 1                   | 3.95                                  | 12.60                           |
| $Z\gamma$  | 005900   | 2.6                                  | 1                   | 3.35                                  | 2.99                            |
| $Zb$       | 005178   | 153.8                                | 0.66                | 205.0                                 | 0.79                            |
| $t\bar{t}$ | 005200   | 449.8                                | 0.73                | 449.8                                 | 0.97                            |

Table 6.3: List of the samples used in 14 TeV analysis, with cross sections  $\sigma$ , NLO/LO  $K$ -factors, average weights  $\langle w \rangle$  related to the NLO, and corresponding integrated luminosities.

## 6.4 Monte Carlo Simulation

The full chain of Monte Carlo generation can be divided into an event generation step and a detector simulation part. The generation of the hard process of interest is based on an analytically calculated matrix element to certain order in perturbation theory. Then follows the fragmentation which simulates the parton emission from the incoming and outgoing partons. In order to obtain a phenomenologically realistic picture partons need to be grouped into hadrons. This process of hadronization is usually done following the string or cluster models. At this point the event is described in terms of physical particles which may undergo various decay processes. Figure 6.16 gives a schematic picture of the event simulation.

The outcome of the simulated collision process is stored in a common format of the HepMC package [123]. This event record is the input to the detector simulation which handles the interaction of particles with matter distributed according to a detailed detector model.

It is challenging to provide a realistic simulation of QCD interactions as they become very complicated when including processes beyond the leading order (LO). Since we are

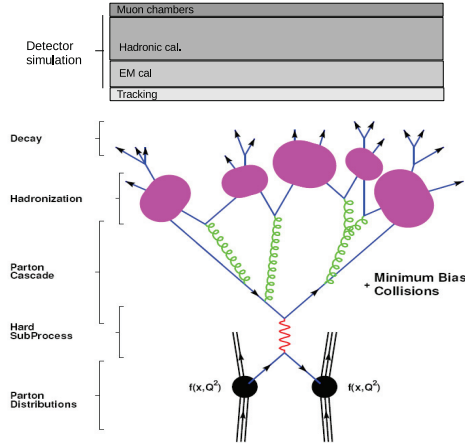


Figure 6.16: Schematic picture of the event simulation (adapted from [123]).

dealing with proton collisions one needs also to simulate the underlying events. ATLAS uses therefore a number of MC generators, often combining a specialized generator for the hard signal process with a general purpose generator for the hadronization and the underlying event.

The generation of the SUSY signal samples starts by a calculation of a consistent set of particle masses and branching fractions for a given mSUGRA parameter set. For this purpose ATLAS uses the ISASUSY [139] component of ISAJET [140]. The next-to-leading-order (NLO) cross sections are obtained via K-factors calculated using PROSPHINO [146]. The event simulation is completed with the HERWIG [141] general purpose generator applying the ISAWIG [147] interface, while the underlying event is simulated using Jimmy [142].

The  $t\bar{t}$  production was simulated using MC@NLO [148, 149] which provides a full next-to-leading-order calculation, while the full event was simulated by a combination of HERWIG and Jimmy. For 10 TeV simulation the MC@NLO is also used to simulate diboson events with subsequent leptonic decay. The diboson samples generated at 14 TeV listed in Table 6.3 were generated using HERWIG only while the NLO cross section was obtained via K-factors. The choice of HERWIG was motivated by the possibility to include off-shell bosons [150].

The AcerMC [151] generator specializes in simulation of processes involving  $W$  and  $Z$  accompanied by jets, including  $b$ -jets. For the 14 TeV simulation it was used to generate the important  $Zb$  background. Among the 10 TeV samples it was also used to generate the single-top events. The simulation of the full event was completed using the general purpose generator PYTHIA [152]. PYTHIA was also used to generate the  $Z\gamma$  and  $W\gamma$  samples, both for 14 and 10 TeV.

Moving to more exact quantum mechanical calculation one encounters interference phenomena whose contributions may have a negative sign. We observe therefore that in the MC@NLO and AcerMC samples there are events with negative weights. This feature has been taken into account when calculating normalization factors.

Processes which involve  $W$  or  $Z$  plus a number of jets represent one of the challenging

SUSY backgrounds. For such events ATLAS is using the specialized ALPGEN [153] generator which is based on calculations of the exact matrix elements for multi-parton processes. In order to obtain a better description of the jet multiplicity, special care is taken to match correctly the jets from the initial matrix-element and jets from parton showering. The latter is provided by the frequently used combination of HERWIG and Jimmy. Although such samples were available for the 14 TeV simulation they could not be used in the trilepton analysis due to a filter at the event generation level which required at least four jets with  $p_T$  above 40 GeV. These cuts are incompatible with the trilepton event selection and the samples could not be used. For the 10 TeV simulation the jet filter was relaxed making it possible to include the ALPGEN samples both for  $W$  and  $Z$  plus  $N$  jets as well as dedicated samples for  $Wbb$  and  $Zbb$  plus  $N$  jets. In addition there are also ALPGEN samples covering QCD jets associated with  $b$ -production. The full list of the 10 TeV samples can be found in Table C.1 in Appendix C and a summary in Section 9.2, Table 9.3.

The tau decays were in most samples treated by the specialized Tauola [154] package. The PHOTOS [155] QED radiation package was used to simulate the electromagnetic radiation from charged leptons.

Both the 14 and 10 TeV samples contain fully simulated events obtained through the production chain described in Section 5.7.1. Full simulation means that the detector response is handled by the detailed GEANT4 [124] detector and particle interaction package.

The last part of the analysis (Chapter 10) uses also samples simulated using the ATLFAST II [156] package. It has two components: Fast ATLAS Tracking Simulation (Fstras [157]) for the inner detector and the muon system and Fast Calorimeter Simulation (FastCaloSim [158]). The speed up is significant as compared to the full GEANT4 simulation, while at the same time, ATLFAST II uses standard ATLAS reconstruction.

## 6.5 Analysis data preparation

The starting point of the analysis has been the Analysis Object Data files (AODs). For the CSC analysis the Athena EventView [159] framework was used in order to obtain small ntuples suitable for quick analysis. Requiring that an event contains leptons highly reduces the size of the samples. The TopView [160] package provided very flexible and transparent framework for producing customized ntuples. Most of the built in top analysis tools were not used during the ntuple production, while additional user-defined truth information and variables were introduced.

Over time ATLAS analysis policies and tools evolved and the analysis of the 10 TeV samples is based on Derived Physics Data (DPD) files. These are a result of skimming, slimming and thinning procedures suitable for a trilepton analysis. The derived files keep the container structure and can be considered as so-called user defined D2PDs.

Both ntuples and DPDs were made utilising the Ganga [102] Grid job management tool and submitted to NorduGrid with the GangaNG backend. NorduGrid handled the production in a very efficient way. Even large data sets like  $t\bar{t}$  with several hundreds of input files were split over 20-200 jobs and analyzed within hours.



# Chapter 7

## Physics Object reconstruction and particle identification performance

Section 4.2 described the functionality of the different parts of the ATLAS detector. When it comes to reconstruction of the particles, one needs to combine information from different detector subsystems. Sophisticated algorithms are implemented in order to obtain precise measurements of track parameters, momenta, energies and other variables. As particles interact in characteristic ways, these patterns are exploited in order to provide as reliable identification as possible. This chapter presents the physics objects used in the analysis, focusing mainly on leptons and isolation requirements.

As jets and missing energy will play an important role in the analysis, these objects will first be briefly introduced.

### 7.1 Jets and missing energy

ATLAS has implemented a number of jet finding algorithms which include both the seeded fixed-size cone algorithms which are not infrared and collinear safe<sup>1</sup> and more sophisticated ones which are safe. Before running the jet finding algorithm on data, the signal from about 200000 calorimeter cells has to be organized into physically meaningful objects. One strategy is to create so-called calorimeter towers which sum the cell energy as it is projected on a fixed grid of  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ . This grid provides a fine granularity, but the procedure does not include any noise suppression or calibration. Another type of ordering strategy is based on so-called topological cell clusters which do not have any fixed size, but rather try to reconstruct three-dimensional “energy-blobs”. This approach involves also calibration and noise suppression.

This analysis uses the fixed cone jet approach. The process of reconstructing jets starts by ordering all input objects according to decreasing transverse momenta. The object with the highest  $p_T$  above a threshold of 1 GeV becomes a cone seed and all objects within the fixed cone of certain size ( $\Delta R$ ) are added to the seed. A new cone is created around the four momentum direction of the resulting object and the summation is repeated until a stable direction is obtained. Once the jets are created a split and merge step is done in order to avoid that objects are included in several jets.

---

<sup>1</sup>A jet algorithm is not infrared and collinear safe if the number of jets provided by the algorithm is ambiguous in the case of soft and collinear parton emission.

Missing transverse energy  $\cancel{E}_T$  is one of the key signatures of many new search channels. It is mainly based on the signal from calorimeters and the reconstructed muon tracks. In addition to the process of interest the analysis of ATLAS data will be affected by a number of background processes like the underlying event, pile-up, and coherent electronics noise which need to be distinguished in the overall energy balance. There are also many sources of fake  $\cancel{E}_T$  for example dead regions and channels, limited detector coverage, finite detector resolution and different kinds of noise.

ATLAS uses two  $\cancel{E}_T$  reconstruction algorithms. The object-based algorithm starts from the reconstructed and calibrated high level objects like electrons, jets and muons. The cell-based algorithm, on the other hand, sums the energy in all calorimeter cells above a noise threshold and corrects it for muon energy and the energy loss in the cryostats. This method is considered to be relatively reliable even during early data taking as it is independent of the reconstruction of other objects. As a final step, the initial global calibration of the calorimeter cells is replaced by a refined one based on the more precise calibration of the reconstructed objects to which these cells are associated. The  $\cancel{E}_T$  variable used in this analysis is a result of this refined algorithm and is referred to as **MET\_RefFinal**.

The expression for the  $\cancel{E}_T$  resolution is  $\sigma = a\sqrt{\Sigma\cancel{E}_T}$  where the parameter  $a$  is between 0.53 and 0.57 for  $\Sigma\cancel{E}_T$  in the range of 20-2000 GeV. The two algorithms show similar performance.

## 7.2 Muon isolation and identification study

Muons are stable particles within the detector and as minimum ionizing particles in the momentum range of interest they have a large penetration power. Muons with  $p_T$  above 3 GeV are able to reach the outer layers of the detector and their main signature is a track through the whole detector volume. Tracks are based on the ionization left by charged particles in the sensitive inner detector layers as well as in the different active parts of the Muon Spectrometer (MS). These so-called “hits” or “space points” are combined by various algorithms into physical particle trajectories. Tracking volumes are immersed in a magnetic field and the curvature provides information both about the charge and the momentum. Since the tracks become less curved at large momenta the resolution deteriorates as the momentum grows. The resolution  $\sigma(p_T)/p_T$  is Gaussian in  $1/p_T$  and it grows linearly with  $p_T$ . The contribution from multiple scattering is close to stable over the whole momentum range.

The muon spectrometer resolution at low energies is dominated by the energy loss in the significant amount of material in front of the MS. For intermediate energies multiple scattering contributes most while at high energies, above  $p_T = 300$  GeV the limiting factors are detector characteristics, alignment and calibration.

Also the calorimeters provide an important support of the muon identification as the muons leave there the characteristic Minimum Ionising Particle (MIP) signature. The Inner Detector provides an independent precise track and momentum measurement which combined with the MS measurement for momenta  $30 < p_T < 200$  GeV gives a better result than the standalone ones.

There are three categories of reconstructed muons [161]:

- *standalone* is a MS-track extrapolated back to the beam line,

- *tagged* is an ID-track, extrapolated to the spectrometer and matched with nearby hits
- *combined* is a standalone muon matched to the closest ID track.

There is a number of muon reconstruction algorithms specialised for the different detector parts. Combined muons are a result of an association of results from algorithms specialized for the ID and for the MS. In ATLAS there are two combined algorithms: Muid [162] and Staco [163]. The latter, which following the recommendation from the muon performance group is used in this analysis, applies a statistical combination of the two independent measurements by means of their covariance matrices. A measure of the quality of the match is given by the  $\chi^2$ , defined as the difference between the squared outer and inner track vectors weighted by their combined covariance matrices.

In simulations performed with ATLAS release 12, it was recommended by the muon performance group to use the Staco algorithm. The same was used during the analysis of the 10 TeV MC samples done with a later release.

Objects are identified as muons if they pass the following requirements:

- Kinematical cuts:  $p_T > 10$  GeV and  $|\eta| < 2.5$
- **bestMatch**: Among the many possible matches between the MS and the ID-tracks the best match is chosen. The  $\chi^2$  of the match is required to be less than 100.
- **isHighPt**: Only muons reconstructed by the algorithm dedicated for high  $p_T$  muons are considered.
- Isolation requirement: various isolation requirements were studied and will be described for each case.

Muons were selected first and were not subject to any overlap removal.

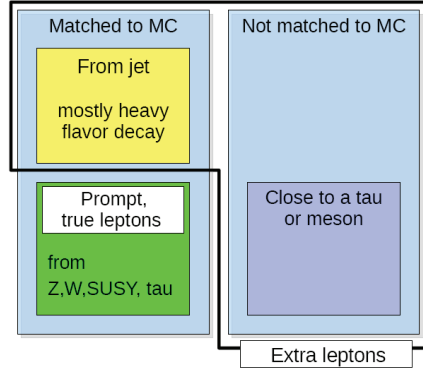
### 7.2.1 Muon MC-truth information

In order to obtain MC information, reconstructed muons are matched to generator level muons, often called “truth muons”, requiring the angular distance  $\Delta R$  (Eq. 4.4) between these two to be less than 0.02. A muon without such a match is in this analysis referred to as *not-matched*. In case of an MC match, an analysis of the chain of mothers have been applied assigning additional flags and adding the mothers to user data.

In summary, we divide the objects that are reconstructed as muons in the following classes <sup>2</sup>:

- *prompt*: A generator level muon is found within  $\Delta R < 0.02$  having one of the following as mother: a SUSY particle, Z, W or  $\tau$
- *from jet*: A generator level muon is found within a  $\Delta R < 0.02$ , but the mother is not one of the above listed particles.
- *not-matched*: No generator level muon is found within  $\Delta R < 0.02$  of the reconstructed muon.
- *extra*: A broader class consisting of muons labelled as *from jet* and *not-matched*. These can be also considered as *non-prompt*.

Figure 7.1: Schematic overview of the different classes of leptons and their labels. The sizes of the boxes do not reflect the fractions these represent in the concrete samples.



A schematic summary of the different classes of leptons is shown in Figure 7.1.

*Extra* muons originate dominantly from jets with a semileptonic heavy flavor decay. In the sample of muons labelled *from jet*, these account for approximately 93% and 83% in the  $t\bar{t}$  and SU3 and other SUSY samples, respectively. Possibly hard gluons accompanying the process under study can split to a  $b\bar{b}$  pair where each  $b$  can in turn produce a muon through weak decays. Other sources of non-prompt muons are meson decays in flight and punch-through from the hadronic calorimeter into the muon chambers.

A match to a generator level muon is needed in order to find the mother particle. However, for a reconstructed muon without such MC counterpart one can look for other generator level particles in its vicinity in terms of angular distance in order to obtain more information about its origin. The analysis ntuples contain a record of all unstable particles in the event, like pions and kaons, and one can look if such particles can be found within a small  $\Delta R$  distance. An attempt was therefore made to match such muons to a generator level tau. A distance of  $\Delta R < 0.1$ , which is the standard requirement for a tau truth matching, was accepted as successful and was found for 4% of the muons. A further search for unstable particles close to these muons showed that more than 80% could be associated to pions, and about 10% to kaons with  $\Delta R < 0.1$ .

The vast majority (95% in SU3, 80% in  $t\bar{t}$ ) of muons without a match to a tau was found to be close to ( $\Delta R < 0.1$ ) one or more unstable particles in the MC record of the event. In samples with a high number of b-jets like  $t\bar{t}$ , could more than 50 % be matched to a B- or D-meson. The lighter flavors were dominant in SU3.

This study showed that the muons without a MC match, and thus without precise knowledge of the mother, originate with high probability from jets. It is therefore convenient to use the notion of *extra* muons referring to both muons *from jet* and those *not-matched* to generator level muons.

Figure 7.2 shows a stack plot of the muon  $p_T$  distributions for the SU3 and  $t\bar{t}$  samples. One can see that the muons from jets and those *not-matched* with a generator level muon, make up a large fraction of the total distribution. Since no isolation cut was yet applied in Figure 7.2, a large number of muons from jets is expected, and the distributions show the characteristic low  $p_T$  of these muons.

<sup>2</sup>The same definitions are applied for electrons.

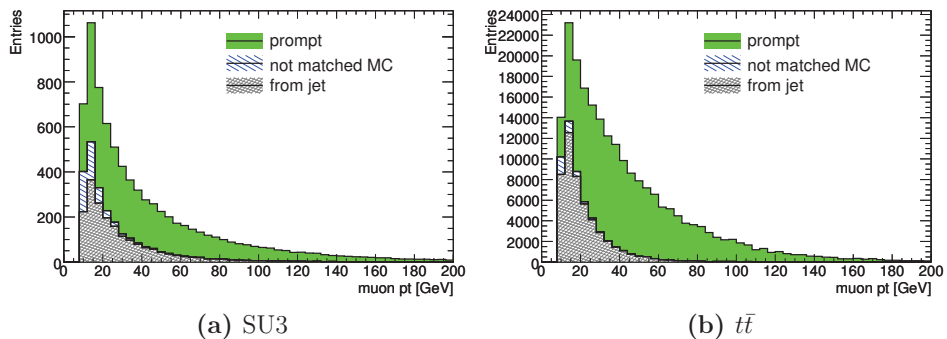


Figure 7.2: Stack plot of  $p_T$  distributions for muons from SU3 (a) and  $t\bar{t}$  (b). No isolation cut is applied. The distributions are normalized to  $1 \text{ fb}^{-1}$ .

## 7.2.2 Muon isolation study

The aim of the isolation cut is to distinguish between prompt signal leptons and *extra* leptons. Hence the representative SU3 SUSY signal sample and the  $t\bar{t}$  sample, which is one of the dominant backgrounds due to presence of *extra* leptons, were selected for a closer study.

Two types of isolation variables have been studied. One is calorimeter-based and it is defined as the sum of the transverse energy deposited in a cone with half opening angle  $\Delta R$  minus the energy in a narrow central part, assumed to represent the particle itself. In many cases one can therefore observe negative values of the isolation energy. As much as approximately 45% of the muons have negative values when the isolation is calculated in the narrow cone with  $\Delta R = 0.2$ . The fraction decreases for wider cones and is reduced to about 23% for the  $\Delta R = 0.45$  cone. The fraction of electrons having negative values of calorimeter-based isolation is approximately 5% for all studied cone widths ( $\Delta R = 0.2, 0.3, 0.45$ ).

This calorimeter-based isolation variable will be denoted by  $E_{\text{cal}}^{\Delta R=0.X}$  where  $0.X$  is the half opening angle of the isolation cone.

A lepton associated with a jet is typically accompanied by a number of charged particles. The second type of isolation information is therefore based on tracks in a cone around the lepton. The track of the lepton itself is not considered. Two track-based isolation variables have been investigated:

- $p_{T\Sigma\text{track}}^{\Delta R=0.X}$  is a sum of  $p_T$  of tracks within a cone of  $\Delta R = 0.X$  and with  $p_T > 1 \text{ GeV}$
- $p_{T\text{max track}}^{\Delta R=0.X}$  [164] is the  $p_T$  of the leading track in a cone with  $\Delta R = 0.X$ . Only tracks with  $p_T > 1 \text{ GeV}$  are taken into account

We start by looking at the calorimeter-based isolation  $E_{\text{cal}}^{\Delta R=0.X}$  as it is a very common choice. The narrow cone of  $\Delta R = 0.2$  is favourable if one wants to ensure a high efficiency for prompt muons. However, as can be seen in Figure 7.3 (a) and (d), the distribution for muons *from jets* and *not matched* to generator level muons stretches also more towards low values. The overlap with the peak of the distribution for prompt muons seems to be larger for the narrow cone than for the wider cones  $\Delta R = 0.3$  and  $\Delta R = 0.45$  shown in Figure 7.3 (b)-(e) and (c)-(f), respectively.

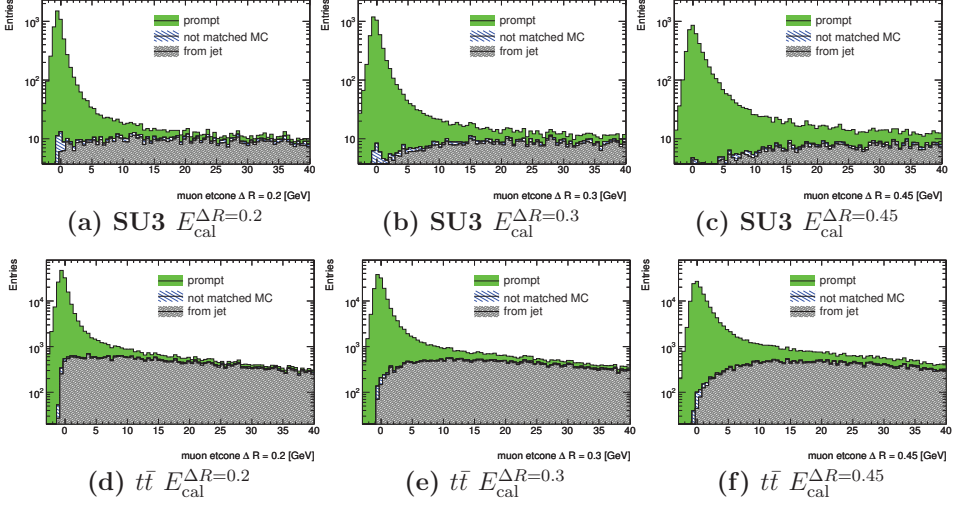


Figure 7.3: Calorimeter-based isolation  $E_{\text{cal}}^{\Delta R=0.X}$  for muons in the signal sample SU3 and background  $t\bar{t}$ . The distributions are normalized to  $1 \text{ fb}^{-1}$ .

The most common jet reconstruction algorithm uses a cone radius of 0.4 and it is perhaps not surprising that narrower isolation cones, like  $\Delta R = 0.2$ , give a considerable probability for accepting locally isolated muons within a jet.

A wider isolation cone like for example  $\Delta R = 0.3$ , Figure 7.3 (b),(e), picks up more of the jet constituents surrounding the lepton leading to a larger difference between isolated prompt muons and muons *from jets*. On the other hand, going to an even larger opening angle like  $\Delta R = 0.45$ , Figure 7.3 (c),(f), a larger fraction of the prompt muons have high isolation energy as the wider cone includes energy from other parts of the event deposited in their vicinity. The same is observed for the track-based isolation variables shown in Figure 7.4.

Since the energy deposited around a lepton may increase for leptons with high momenta, specially in the case of electrons, it is common to construct a so-called relative isolation. The latter compensates for the  $p_T$  dependence by dividing the isolation by the  $p_T$  of the particle. The relative track-based isolation for muons is shown in Figure 7.5.

Prompt signal leptons from SUSY represented by SU3 and *extra* leptons from the dominant background  $t\bar{t}$  were selected for a closer study of the isolation  $p_T$  dependence. Figure 7.6 shows the profile plot of the  $E_{\text{cal}}^{\Delta R=0.3}$  and  $p_{T\Sigma\text{track}}^{\Delta R=0.3}$  vs.  $p_T$  for these two classes of muons. The distribution is nearly flat for prompt muons, something that can be explained by the fact that muons have a small probability for emitting Bremsstrahlung and deposit very little energy in the calorimeters, and are therefore less likely to contribute to the isolation energy summed up in the cone. On the other hand, for *extra* muons from  $t\bar{t}$  we see a clear positive slope.

The profile plot of the relative isolation as function of  $p_T$  is shown in Figure 7.7. The distribution for *extra* muons has a clear negative slope and is well above the profile plot for the prompt muons, specially in the low- $p_T$  region, where most of the *extra* muons are found.

This section discussed calorimeter  $E_{\text{cal}}^{\Delta R=0.X}$  and track  $p_{T\Sigma\text{track}}^{\Delta R=0.X}$  based isolation variables

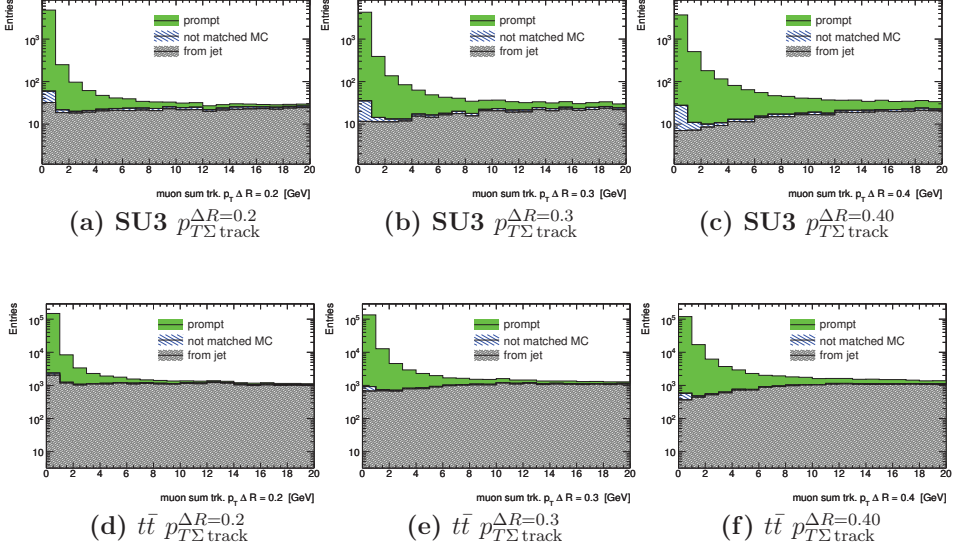


Figure 7.4: Track-based isolation  $p_{T\Sigma}^{\Delta R=0.X}$  for muons in the signal sample SU3 and background  $t\bar{t}$ . The distributions are normalized to  $1 \text{ fb}^{-1}$ .

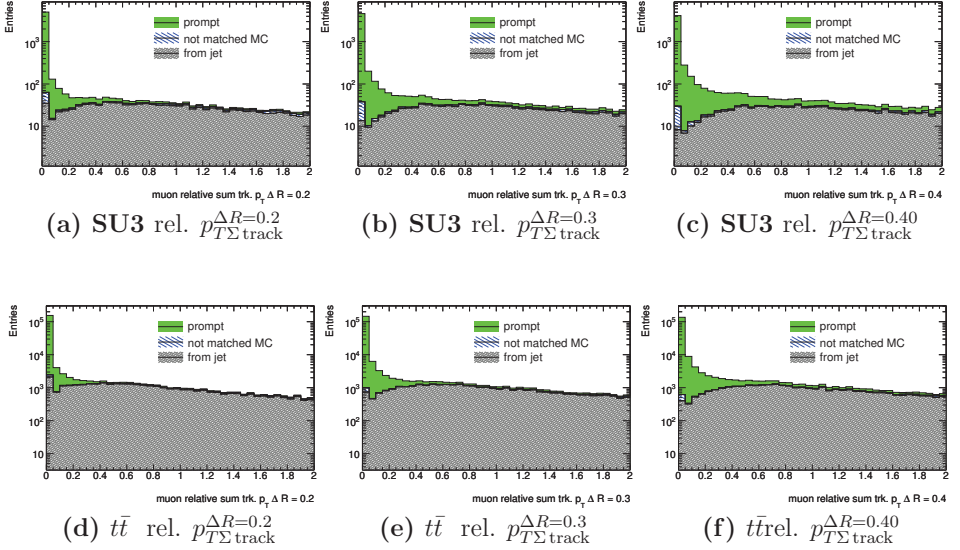


Figure 7.5: Relative track-based isolation  $p_{T\Sigma}^{\Delta R=0.X}$  for muons in the signal sample SU3 and background  $t\bar{t}$ . The distributions are normalized to  $1 \text{ fb}^{-1}$ .

as well as their  $p_T$  dependencies for various classes of muons. The next section will deal with the muon-jet separation requirement.



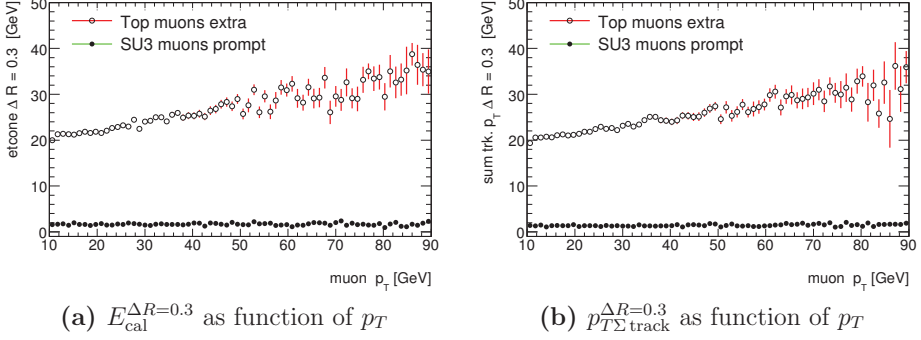


Figure 7.6: Profile plot of  $E_{\text{cal}}^{\Delta R=0.3}$  and  $p_{T\Sigma\text{track}}^{\Delta R=0.3}$  as function of  $p_T$  for prompt muons in SU3 and for *extra* muons in the  $t\bar{t}$  background sample.

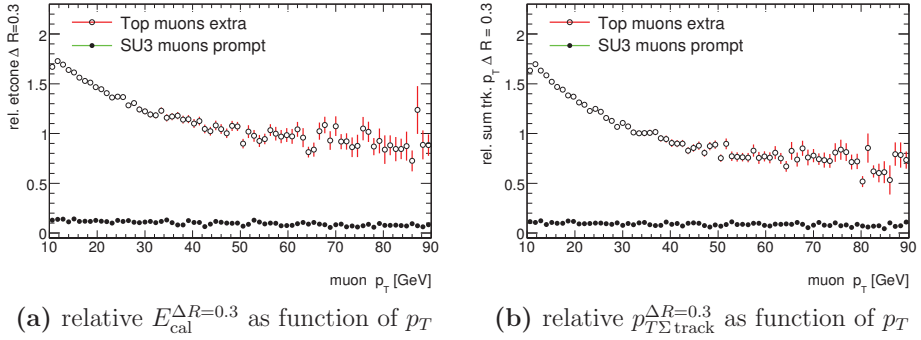


Figure 7.7: Profile plot of  $E_{\text{cal}}^{\Delta R=0.3}$  and  $p_{T\Sigma\text{track}}^{\Delta R=0.3}$  as function of  $p_T$  for prompt muons in SU3 and for *extra* muons in the  $t\bar{t}$  background sample.

### 7.2.3 Muon-jet separation

A muon closer to the jet axis than  $\Delta R = 0.4$  is according to the jet cone algorithm inside a jet and is with high probability non-isolated. A generic isolation cut<sup>3</sup>,  $E_{\text{cal}}^{\Delta R=0.2} < 6$  GeV applied on the muons which did not pass the  $\Delta R(\mu, \text{jet}) > 0.4$  requirement showed that 90%, 95 % of the *extra* (non-prompt) muons in  $t\bar{t}$ , SU3 respectively, could be considered as non-isolated.

A strong rejection of *extra* muons inside jets yields a significant reduction of the total number of *extra* muons in a sample since approximately 94 % of all such muons are found inside jets, both in SU3 and  $t\bar{t}$ <sup>4</sup>. The distance to the closest jet can, in addition to the isolation cut, be exploited as an efficient tool in rejecting *extra* muons. A jet can contain locally isolated muons since the isolation cone is narrower than the jet cone. Figure 7.8 shows that the muon jet separation requirement can remove the vast majority of isolated *extra* muons (marked by solid squares). On the other hand, Figure 7.8 shows that there is also a large fraction of prompt muons that do not pass the muon-jet separation requirement, 24 %, 19 % and 24 % in SU3, SU2 and SU4, respectively. The majority of

<sup>3</sup>This is not a cut optimized for trilepton analysis, it is chosen to illustrate the point. This cut is often used in  $t\bar{t}$  analysis.

<sup>4</sup>In the  $Zb$  sample 77 % of *extra* muons did not pass the separation cut  $\Delta R(\mu, \text{jet}) > 0.4$ .



these is actually well isolated and has a relatively high momentum.

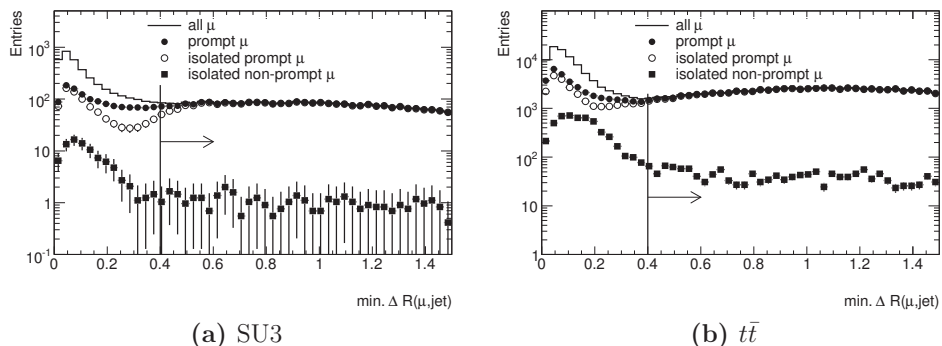


Figure 7.8:  $\Delta R$  between muons and jets (jets have been through overlap removal with electrons). Histograms show muons in SU3 and  $t\bar{t}$  samples. The distributions for other SUSY samples are similar.

Figure 7.9 (a) shows the profile plot of the minimum angular distance between muons and jets,  $\Delta R_{\min}(\mu, jet)$ , after applying the test isolation criterion  $E_{\text{cal}}^{\Delta R=0.2} < 6$  GeV. Both in SUSY and  $t\bar{t}$  the jet density is very high, specially in the central barrel region leading to a V-shaped distribution in Figure 7.9 (a). At  $|\eta| \sim 1$  one can see a small bump on top of the V-form. This  $\eta$  region corresponds to the transition between the barrel and the extended hadronic calorimeter, where jet reconstruction is more difficult. The complex detector conditions make also the cut on the hadronic leakages a less efficient tool for discriminating between jets and electrons. This leads to a higher electron fake rate, which means that a larger number of jets is wrongly identified as electrons and thus not considered when calculating  $\Delta R_{\min}(\mu, jet)$ .

Figure 7.9 (b) shows the  $p_T$  dependence of  $\Delta R_{\min}(\mu, jet)$ . The mean minimum distance between muons and jets is decreasing with growing muon  $p_T$ . This may be related to the fact that parents of the signal muons in long SUSY decay chains may be highly boosted and thus closer to high  $p_T$  jets from the same decay chain. Because of this effect, prompt high  $p_T$  muons will have a smaller probability to pass the muon-jet separation requirement and will have a significantly decreasing selection efficiency for high  $p_T$  muons as shown in the trilepton analysis in [78] and as Figure 7.22 (Section 7.5) will confirm.

Aiming at high muon selection efficiency, one can question whether this is an optimal cut in a jet-rich environment like we find in SUSY events and given the high lepton-momenta and boosts. Instead of simply rejecting all muons regarded as inside a jet, the idea is to make them subject to strict additional cuts on typical features of *extra* muons. The potential of three different cuts have been studied and discriminated between:

- First alternative: Figure 7.10 (a) shows the transverse impact parameter significance  $|d_0/\sigma(d_0)|$  along the x-axis and the relative track-based isolation  $p_{T\Sigma\text{track}}^{\Delta R=0.2}$  along the y-axis. An isolation cut of  $< 0.1$  or so, corresponding to a horizontal line, would remove a large fraction of the *extra* muons. An *extra* cut on  $|d_0/\sigma(d_0)|$  could possibly give an even cleaner muon sample. However, it can be questioned if the cut is efficient enough. A study of the  $|d_0/\sigma(d_0)|$  distributions for prompt muons in SUSY and *extra* muons from the background samples  $t\bar{t}$  and  $Zb$  shows that the mean values are the same and only the width is different, making the results in general not promising.

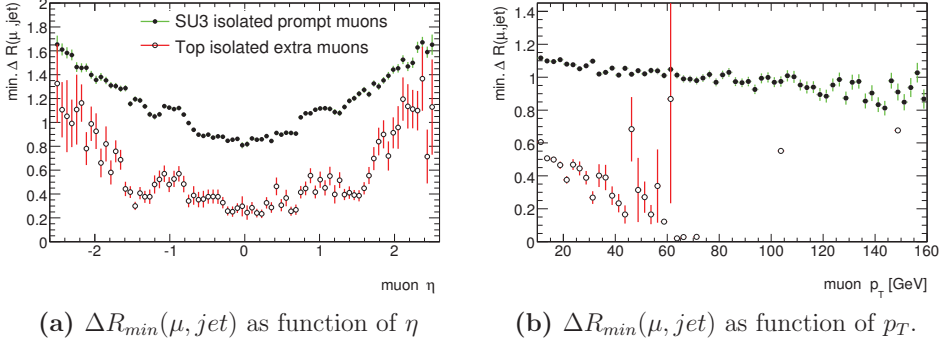


Figure 7.9: Profile of  $\Delta R$  angular distance to closest jet for isolated prompt muons in SU3 and isolated *extra* muons in  $t\bar{t}$ . The isolation cut is  $E_{cal}^{\Delta R=0.2} < 6$  GeV.

- Second alternative: The two most promising isolation variables, the relative calorimeter-based isolation and the relative track-based isolation, were combined by a linear Fisher discriminant. The method performs well, although it needs careful training and the variables are not really Gaussian.
- Third alternative: It has been shown in Section 7.2.2 that the relative track-based isolation performs well for low  $p_T$  providing good separation between *prompt* and *extra* muons. Figure 7.10 (b) shows a scatter plot of the  $p_T$  along the x-axis and relative  $p_{T\Sigma track}^{\Delta R=0.2}$  along the y-axis for muons within a jet,  $\Delta R(\mu, jet) < 0.4$ ; for *prompt* muons from SU3 and *extra* muons from  $t\bar{t}$ . A linear cut as indicated in the plot combines a higher  $p_T$  threshold with an isolation cut. The slope of the line has been suggested by combining these two variables in a Fisher discriminant, while the constant terms are chosen such that they assure a minimum  $p_T$  of 30 GeV or 40 GeV. As shown in Table 7.1, Row 3 and 4, the relative  $p_{T\Sigma track}^{\Delta R=0.2}$  gives better results than the relative  $p_{T\Sigma track}^{\Delta R=0.3}$ . The lines giving best results were:

$$f_{30}(p_T) = 0.003(p_T - 30 \text{ GeV}) \quad (7.1)$$

$$f_{40}(p_T) = 0.003(p_T - 40 \text{ GeV}) \quad (7.2)$$

The subscript indicates the  $p_T$  threshold. These two lines are indicated on the scatter plot in Figure 7.10 (b). Muons below the line are accepted.

Such a linear cut is relatively simple and robust. Compared to the Fisher discriminant, it is less sensitive to many factors that could be uncertain during early data taking like for example the shape of the MC distributions. At the same time, it gives high rejection of *extra* muons.

A simple qualitative optimization study was performed. A cut which gave a selection efficiency for *extra* muons from  $t\bar{t}$  lower than 1% was applied and the corresponding efficiencies for prompt muons from the signal sample SU3 are compared. Table 7.1 shows the fraction of muons that pass the cut. The Fisher discriminant combining the relative  $p_{T\Sigma track}^{\Delta R=0.2}$  and  $E_{cal}^{\Delta R=0.2}$  is possibly slightly better than a combination of the same variables calculated in the  $\Delta R = 0.3$  cone. The linear cut proposed as the third alternative shows

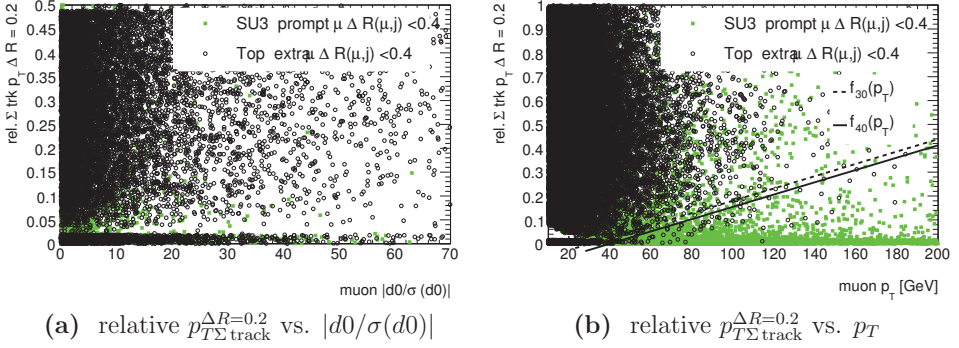


Figure 7.10: Scatter plot of the transverse impact parameter significance  $d0/\sigma(d0)$  vs. relative track-based isolation  $p_{T\Sigma \text{ track}}^{\Delta R=0.2}$  along the y-axis (a).  $p_T$  along the x-axis,  $p_{T\Sigma \text{ track}}^{\Delta R=0.2}$  along the y-axis (b). The plots show muons inside a jet with  $\Delta R(\mu, jet) < 0.4$ . The lines in (b) indicate the linear cuts  $f_{30}(p_T)$  given by Eq. 7.1 and  $f_{40}(p_T)$  by Eq. 7.2.

also very good performance. The last line in the table shows the results for the linear cut given by the  $f_{40}(p_T)$  line and the high  $p_T$  threshold rejects as expected a large fraction of the *extra* muons in  $t\bar{t}$  events. Only 0.2% pass the cut. The efficiency for prompt muons is also reduced, however, knowing that the  $t\bar{t}$  production has a much higher cross section than the signal, it is more important to aim at good rejection of *extra* muons.

|   | SU3 <i>prompt</i> $\mu$ | SU3 <i>extra</i> $\mu$ | $t\bar{t}$ <i>prompt</i> $\mu$ | $t\bar{t}$ <i>extra</i> $\mu$ |
|---|-------------------------|------------------------|--------------------------------|-------------------------------|
| Fisher $E_{\text{cal}}^{\Delta R=0.3}$ & $p_{T\Sigma \text{ track}}^{\Delta R=0.3}$ | 57%                     | 0.3%                   | 63%                            | 0.7%                          |
| Fisher $E_{\text{cal}}^{\Delta R=0.2}$ & $p_{T\Sigma \text{ track}}^{\Delta R=0.2}$ | 59%                     | 0.2%                   | 59%                            | 0.4%                          |
| $p_T$ vs. rel. $p_{T\Sigma \text{ track}}^{\Delta R=0.3}$ $p_T > 30\text{GeV}$      | 55%                     | 2.0%                   | 59%                            | 0.6%                          |
| $p_T$ vs. rel. $p_{T\Sigma \text{ track}}^{\Delta R=0.2}$ $p_T > 30\text{GeV}$      | 58%                     | 0.6%                   | 63%                            | 0.4%                          |
| $p_T$ vs. rel. $p_{T\Sigma \text{ track}}^{\Delta R=0.2}$ $p_T > 40\text{GeV}$      | 53%                     | 0.5%                   | 51%                            | 0.2%                          |

Table 7.1: The table shows the fraction of muons in a jet ( $\Delta R(\mu, jet) < 0.4$ ) that pass the  $p_T$ -dependent isolation cut. The last line representing the  $f_{40}(p_T)$  provides the strongest rejection of *extra* muons in  $t\bar{t}$  events.

These combined cuts were applied only on muons which do not pass the muon-jet separation requirement. The aim was to remove more than 99% of the *extra* muons while possibly keeping some of the prompt muons that happened to be close to a jet. Since 90-95 % of all *extra* muons are within  $\Delta R(\mu, jet) < 0.4$ , this means that such a cut leads to a rather pure muon sample. At the same time one can still keep approximately half of the 23 % (SU3)<sup>5</sup> of prompt muons which are close to a jet. A study of muons inside a jet which satisfy the  $f_{40}(p_T)$  cut has shown that there are very few such *extra* leptons left and only a small fraction of these can be removed by an isolation cut. In order to remove the remaining *extra* muons, basically those that are found outside the jet cone, it is observed that the relative track-based isolation is most efficient.

Figure 7.11 shows the distributions of the relative track-based isolation variable for prompt muons from SU3 and *extra* muons from  $t\bar{t}$ , normalized to 1, after applying the

<sup>5</sup>The fraction is similar in the other SUSY samples, slightly lower in SU2.

$f_{40}(p_T)$ -cut. The plot shows that an isolation cut will remove a large fraction of the *extra* muons. However, one should bear in mind that the plots are normalized to unity and the actual number of remaining *extra* muons is small. An isolation cut will in fact also remove a considerable fraction of the *prompt* muons.

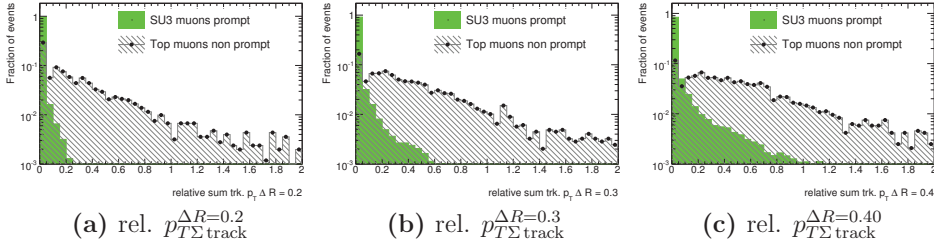


Figure 7.11: Track-based isolation after applying the  $f_{40}(p_T)$ -cut for prompt muons in the signal sample SU3 and muons from jet in  $t\bar{t}$ . All histograms are normalized to 1.

### 7.3 Electron isolation and identification study

Electrons as charged particles leave a track in the inner detector and create a shower in the electromagnetic calorimeter via sequential processes of Bremsstrahlung and pair production. Electron reconstruction which is based on an ID track matched with an EM-calorimeter cluster is optimized to reject fake electrons from QCD jets. In the moderate  $p_T$  range (20-50 GeV) the rejection factor is required to be as high as  $10^5$  in order to obtain a decent signal to background ratio. The jet-rejection factor decreases rapidly with increasing  $p_T$ . Requiring a multi lepton final state suppresses strongly the QCD background introduced due to fake electrons [78].

ATLAS has implemented two complementary electron reconstruction algorithms<sup>6</sup>. The standard algorithm optimized for high  $p_T$  electrons starts with an EM calorimeter cluster reconstructed by the “sliding window” algorithm and matches it with an ID track required to be within a  $\Delta\eta \times \Delta\phi$  window of  $0.05 \times 0.10$ . The ratio of the energy of cluster and momentum of the matched track,  $E/p$  is required to be less than 10. This reconstruction algorithm is referred to as **egamma**. The second algorithm, optimized for soft electrons with energies down to a few GeV, starts with a good-quality track and matches it with an isolated energy deposit in the EM calorimeter. The information about the reconstruction algorithm of an electron can be obtained from the “author” variable.

Electron identification is based on cuts on a number of variables which distinguish electrons from other particles, mainly hadrons. The cuts are grouped in three main quality classes, “loose”, “medium” and “tight”. A loose electron is based on limited information from the calorimeter and satisfies cuts on hadronic leakage and shower shapes described in Table 7.2. The medium cuts use also information from the stripes of the first layer of the EM calorimeter which, with its fine granularity, can reject  $\pi^0 \rightarrow \gamma\gamma$  decays. Cuts on the track quality and the matching of the cluster to a track are also introduced. The best rejection of fake electrons, but at the same time lowest efficiency is obtained with “tight” cuts. This set includes further requirements on the track matching and  $E/p$  and

<sup>6</sup>A third algorithm is dedicated to reconstruction of forward electrons, but these are outside the scope of this analysis.

| Cut description   | Bit name                | Cut                         |
|---|-------------------------|-----------------------------|
| <b>Loose cuts</b>   |                         |                             |
| Acceptance of the detector  | ClusterEtaRange         | $ \eta  < 2.47$             |
| Ratio of $E_T$ in the first sampling of the hadronic calorimeter to $E_T$ of the EM cluster   | ClusterHadronicLeakage  | $< 0.015 - 0.045$           |
| <b>Second layer of the EM calorimeter:</b><br>Ratio in $\eta$ of cell energies in $3 \times 7$ versus $7 \times 7$ cells  | ClusterMiddleEratio37   | $< 0.600 - 0.910$           |
| Lateral width of the EM shower<br>(to reject pions with wide showers)   | ClusterMiddleWidth      | $< 0.014 - 0.020\text{m}$   |
| <b>Medium cuts</b>  |                         |                             |
| <b>First layer of the EM calorimeter:</b><br>Ratio of the energy in the first sampling to the cluster energy  |                         | $< 0.005$                   |
| The crack region $1.37 <  \eta  < 1.52$ is rejected   |                         |                             |
| Second largest energy deposit normalized to the cluster energy  | ClusterStripsDeltaEmax2 | $< 0.25 - 0.52$             |
| Difference between energy associated with the second largest energy deposit and energy associated with the minimal value between the two leading maxima (rejects jets with $\pi^0$ decays that often have two maxima) | ClusterStripsDeltaE     | $< 0.10 - 0.30 \text{ GeV}$ |
| Total shower width  | ClusterStripsWtot       | $< 1.40 - 4.00$             |
| Fraction of energy outside of the core of the EM shower   | ClusterStripsFracm      | $< 0.20 - 0.80$             |
| Width of the 3 strips around the maximum energy deposit   | ClusterStripsWeta1c     | $< 0.6 - 0.8$               |
| <b>Track Quality:</b>   |                         |                             |
| Number of hits in the pixel layer   | TrackPixel              | $\geq 1$                    |
| Number of hits in the pixel and SCT detectors   | TrackSi                 | $\geq 9$                    |
| Transverse impact parameter   | TrackA0                 | $\leq 1\text{mm}$           |
| $\Delta\eta$ between the cluster and the track  | TrackMatchEta           | $\leq 0.005$                |

Table 7.2: List of electron identification cuts applied to obtain a “medium” quality electron. The medium cuts are applied on top of the loose cuts. Some cut values may be slightly different in release 14. Information stems from [78].

use information from the TRT and the b-layer of the Pixel detector. An isolation cut is also available.

Each cut is associated with a bit in the `isEM` bit variable where certain level of quality corresponds to a bit mask. Table 7.2 lists the cuts associated with the “medium” electron masks.

Objects are defined electrons if they pass following cuts:

- Kinematical cuts:  $p_T > 10 \text{ GeV}$  and  $|\eta| < 2.5$
- Further acceptance cut: The isolation energy in the transition region between the barrel and the endcap,  $|\eta| \in [1.37, 1.52]$ , is in ATLAS Software release 12 highly overestimated due to a bug in the code. Events with electron candidates in this region are not considered.
- Quality cuts: electrons are to be reconstructed with the `egamma` algorithm. Electrons reconstructed only by the `soft electron` algorithm are not accepted, which means that the “author” variable is required to be 1 or 3. The `isEM` “medium” quality flag is required.

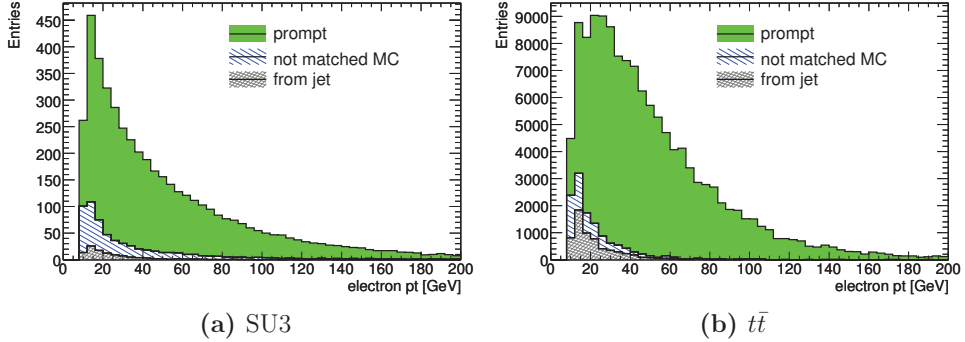


Figure 7.12: Stack plot of  $p_T$  distributions for electrons from SU3 (a) and  $t\bar{t}$  (b). No isolation cut is applied. The distributions are normalized to  $1 \text{ fb}^{-1}$ .

- Isolation requirement: various isolation requirements were studied and will be described for each case as these are the main focus of this study.

In addition to being reconstructed by dedicated electron/photon tools, electrons are also reconstructed by the jet algorithms. Jets are selected after electrons and to avoid including the same physical object both as an electron and a jet, the latter are discarded if the angular distance  $\Delta R$  between the jet axis and an already selected electron is smaller than 0.2. This cut value is based on a study of the  $\Delta R$  between electrons and the closest jet, which could be the electron itself, see Figure 7.18. Jets overlapping with photon and tau objects are kept since the latter are not part of the subject of this analysis and will not be double counted. Jets overlapping with muons are also kept.

After introducing the definition of the electron objects, the next section will discuss in more detail the origin of the *not-matched* electrons.

### 7.3.1 Electron MC-truth information

The method for obtaining MC truth information as explained in Section 7.2.1 for muons is also applied to electrons.

Electrons selected by the procedure above are matched to generator level electrons, often called “truth electrons”, requiring the matching  $\Delta R < 0.02$ . An electron without such a match is in this analysis referred to as *not-matched*. A *prompt* electron is defined as having one of the following particles as mother: a SUSY particle, Z, W or  $\tau$ . However there is also a significant contribution of electrons *from jets*, especially from semi-leptonic b-decays. Before any isolation cut, b-jets account for approximately 96% and 80% of the electrons *from jets* in the  $t\bar{t}$  and SU3 sample, respectively. Other possible sources of *extra* electrons are  $\pi^0 \rightarrow \gamma e^+ e^-$  decays,  $\pi^\pm$ -decay in flight,  $\pi^\pm \pi^0$  combinations, and electrons from photon conversion. The different classes of electrons follow the same definitions as established for muons in Section 7.2.1.

Figure 7.12 shows a stack histogram of the electron  $p_T$  distributions for the SU3 signal sample (a) and  $t\bar{t}$  background (b). The sum of these three histograms represents the full sample. No isolation requirements are applied in these plots. The presence of electrons from jets is clearly more prominent in the  $t\bar{t}$  sample. Figure 7.12 shows that the fraction of *not-matched* electrons is larger in the SU3 than in  $t\bar{t}$ . In order to understand better

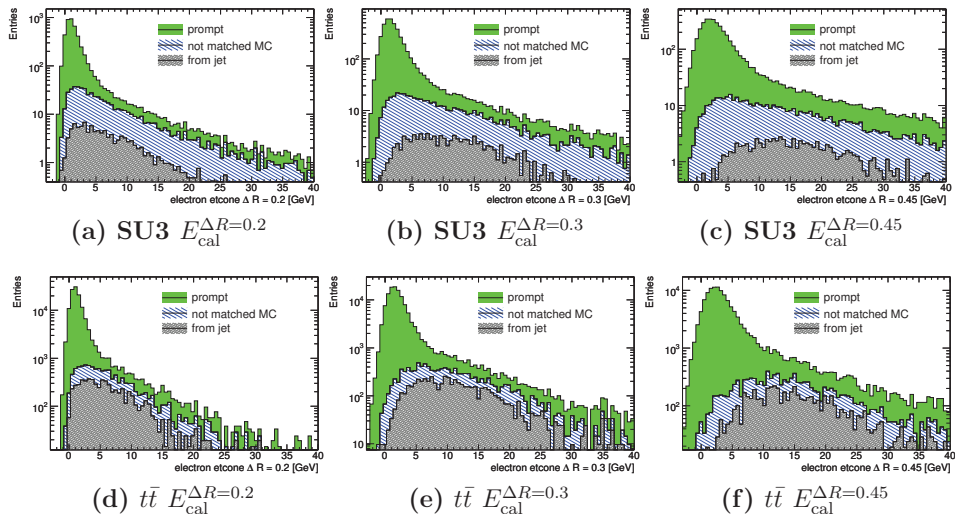


Figure 7.13: Calorimeter-based isolation  $E_{\text{cal}}^{\Delta R=0.X}$  for electrons in the signal sample SU3 and  $t\bar{t}$  background sample. The distributions are normalized to  $1 \text{ fb}^{-1}$ .

the origin of the *not-matched* electrons, one can look for unstable particles within a small  $\Delta R$ .

As a first step, for electrons without a generator level counterpart an attempt was made to find a nearby generator level tau. The electron is considered as associated to a tau if a match is found within  $\Delta R < 0.1$ .

In the tau-rich SUSY sample SU3 approximately 40% of the *not matched*-electrons are found very close to a generator level tau. For the  $t\bar{t}$  sample this fraction is 17%, and 14% for SU1 which has also high tau branching fractions, while for SU2 and SU4 it is down to approximately 2%. A further study of the unstable particles in the event which can be found close to (meaning within  $\Delta R < 0.1$ ) a *not matched*-electron earlier associated to a tau, shows that more than  $\sim 95\%$  could be associated to charged pions and in rarer cases to kaons, which is in good agreement with the tau decay pattern.

The remaining electrons in the *not-matched* class were associated to unstable particles and such was found in  $\sim 70\%$  of cases within  $\Delta R < 0.1$ . Like in the muon analysis, in samples where a high number of b-jets is expected (e.g.  $t\bar{t}$  or SU4) more than 50% of the studied *not-matched* electrons were found to be close to a heavy D- or B-meson, the rest close to light flavor mesons. In SU3 pions accounted for approximately 60% of the associations, while heavy flavor mesons for only a few percent.

One can therefore assume that the majority of the objects reconstructed as *not-matched* electrons can be considered as originating from hadronic jets or from taus. The next section will discuss the isolation of the various types of electrons.

### 7.3.2 Electron isolation study

The isolation variables used for electrons are the same as those defined for muons in Section 7.2.2. However, electrons and muons behave differently in the detector and the mean values of the isolation distributions for electrons have usually higher values.



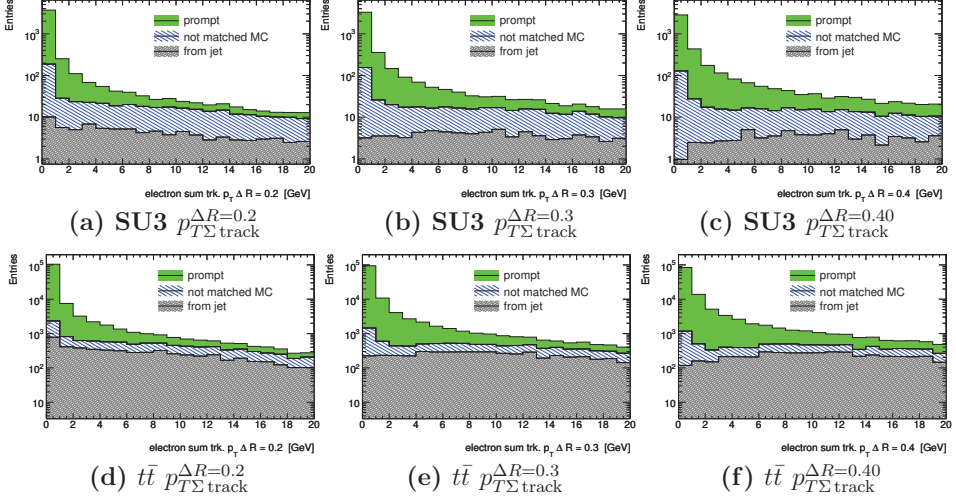


Figure 7.14: Track-based isolation  $p_{T\Sigma}^{\Delta R=0.X}$  for electrons in the signal sample SU3 and background  $t\bar{t}$ . The distributions are normalized to  $1 \text{ fb}^{-1}$ .

Figures 7.13 (a)-(c) (SU3) and (d)-(f) ( $t\bar{t}$ ) show the calorimeter-based isolation  $E_{\text{cal}}^{\Delta R=0.X}$  for different cones:  $\Delta R = 0.2$  and for the wider cones  $\Delta R = 0.3$  and  $\Delta R = 0.45$ .

Due to the same effect as for muons, isolation in the narrow cone, Figure 7.13 plot (a) and (d), is more peaked at low values for the prompt electrons giving a good selection efficiency. However, electrons *from jet* and those without a MC match have also lower isolation energy and the distributions are more signal-like than for the wider cones shown in Figures 7.13 plots (b)-(c) and (e)-(f).

The larger cones may potentially give a better separation between the prompt and *extra* electrons at the price of a larger fraction for prompt electrons picking up more of the surrounding activity, leading thus to higher values of the isolation energy.

The same trend can be observed in the distributions for the track based isolation variables shown in Figure 7.14.

Figures 7.13 and 7.14, plots (a)-(c) show clearly that the SU3 SUSY sample has a large contribution of electrons which are not matched to generator level electron. The earlier described matching to unstable particles in the event showed that a large fraction of the *not-matched* electrons originate from (light)-meson decays, and in the case of SU3 also from  $\tau$ 's, abundantly produced in this parameter space point. These SU3 *not-matched* electrons, of which 40% could be associated to a generator level tau, are more isolated than the *not-matched* electrons in the  $t\bar{t}$  sample where they follow more closely the distribution for electrons *from jet*. For the *not-matched* electrons in SU3, the shift towards higher values of isolation energy for wider cones is smaller as compared to the *not-matched* electrons in  $t\bar{t}$  shown (Figures 7.13 and 7.14, plots (c)-(f)).

Electrons emit easily Bremsstrahlung and this process is known to be momentum-dependent. Figure 7.15, presenting the profile plots of the isolation as a function of  $p_T$ , shows that both the calorimeter- and track-based isolation energy are slightly increasing for higher  $p_T$  and the slope is steeper for *extra* electrons than for prompt electrons. The distribution of  $p_{T\Sigma}^{\Delta R=0.3}$  for prompt electrons, Figure 7.15 (b), is flatter than the calorimeter-based (a). The plots indicate that the track-based isolation is less  $p_T$ -dependent and also



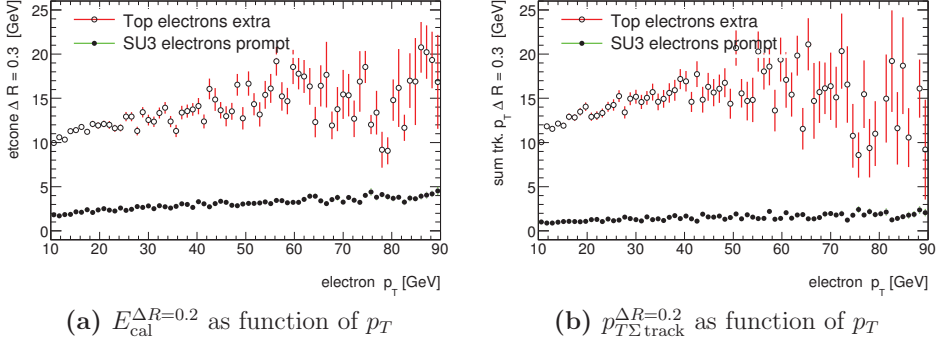


Figure 7.15: Profile plot of  $E_{\text{cal}}^{\Delta R=0.2}$  and  $p_{T\Sigma \text{ track}}^{\Delta R=0.2}$  as function of  $p_T$  for prompt electrons in SU3 and *extra* electrons in the  $t\bar{t}$  background sample.

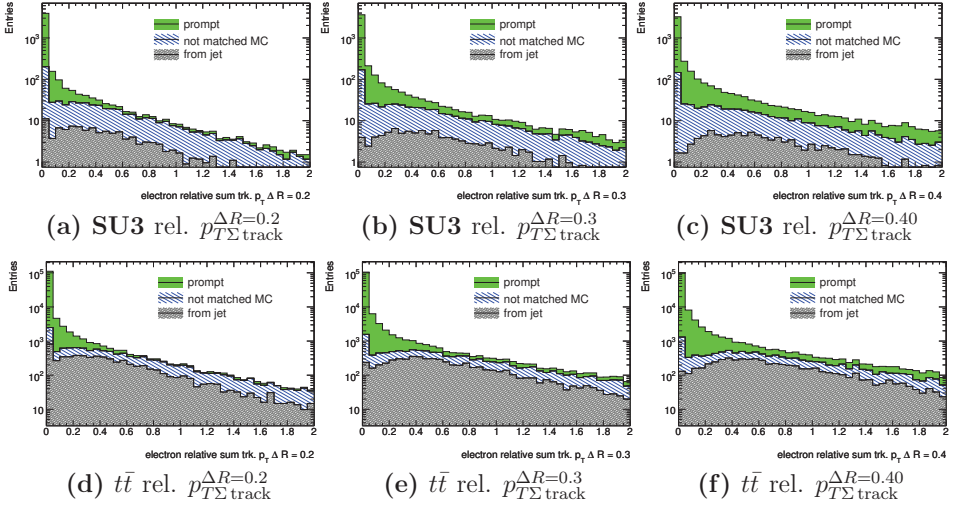


Figure 7.16: Relative track-based isolation  $p_{T\Sigma \text{ track}}^{\Delta R=0.X}$  for electrons in the signal sample SU3 and background  $t\bar{t}$ . The distributions are normalized to  $1 \text{ fb}^{-1}$ .

gives a better separation between the *prompt* and *extra* electrons. The distributions for other cones and samples show the same trends.

A possible way of reducing the  $p_T$  dependence is to use the so-called relative isolation achieved by dividing the isolation energy by the  $p_T$  of the lepton. The distributions of the relative track-based isolation for electrons are shown in Figure 7.16. Compared to Figure 7.14 it seems like there are fewer *prompt* electrons with higher values of relative track-based isolation. The distributions for electrons from jet and those not matched to a MC-truth particles, is almost flat in Figure 7.14, while the relative variable in Figure 7.16 has a clear top shifted away from zero.

Profile plots of the relative electron isolation as function of  $p_T$  are shown in Figure 7.17. The relative track-based isolation, Figure 7.17 (b), for prompt electrons have a weaker  $p_T$  dependence and a more clear separation with respect to *extra* electrons over a larger  $p_T$  range, than the calorimeter-based (a). This separation is enhanced in the low  $p_T$ -region

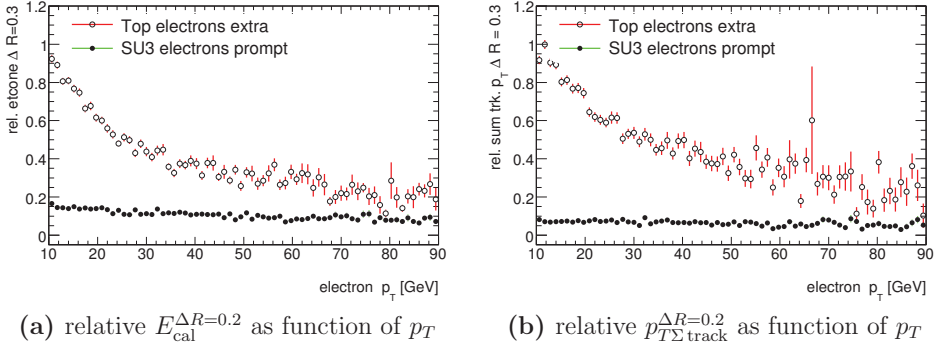


Figure 7.17: Profile plot of relative  $E_{\text{cal}}^{\Delta R=0.2}$  and relative  $p_{T\Sigma \text{ track}}^{\Delta R=0.2}$  as function of  $p_T$  for prompt electrons in SU3 *extra* electrons in the  $t\bar{t}$  background sample.

which is very positive as most of the *extra* electrons have a relatively low  $p_T$ .

It is observed during this isolation study that in a jet, and especially in a  $b$ -jet rich environment, a wider cone with  $\Delta R = 0.3$  may be considered as an alternative to the often used narrow  $\Delta R = 0.2$  cone. The  $p_T$  dependence of the isolation is a strong motivation for using the relative isolation.

### 7.3.3 Electron-jet separation

The procedure of electron-jet overlap removal is applied as described in Section 7.3. Figure 7.18 (a) shows the minimum angular distance between electrons and the closest jet when the overlap removal is not applied and it is clear that the large majority of prompt electrons can be matched to their corresponding jet object within  $\Delta R = 0.2$ , which is also the cut for overlap removal. The solid line histogram shows the  $\Delta R$  to all jets in the event, not only the closest. Therefore we can see a long tail for high  $\Delta R$  which represents matches to more distant jets. It is observed that very few electrons find a "closest jet" beyond  $\Delta R = 0.4$ . It is also interesting to note that the mean of the distribution of angular distance for isolated *non-prompt* electrons (solid squares in Figure 7.18 (a)) is 0.082, which is larger than 0.057 obtained for prompt electrons. This slight difference is expected for electrons which are jet constituents. The isolation requirement used in this plot was  $E_{\text{cal}}^{\Delta R=0.2} < 6$  GeV. Figure 7.18 (b) shows the same distributions after overlap removal. Now the jets which overlapped with a selected electron are rejected and the high values of minimum  $\Delta R$  indicate that the closest jet is no longer the electron itself but some other jet in the event.

The jet algorithm used in this analysis is based on a cone with  $\Delta R = 0.4$  which raises the question how to treat the electrons that match a jet with  $0.2 < \Delta R < 0.4$ . The isolation cut should in principle reject electrons in jets, however it is observed that locally isolated electrons can be found within this intermediate distance to the jet axis which, according to the jet algorithm, is still inside its cone. Requiring the angular separation between an electron and a jet to be larger than 0.4 could reduce the number of *non prompt* electrons. The cut, referred to as  $\Delta R(e, \text{jet}) > 0.4$ , has been applied in the CSC book [78] in order to obtain as pure electron sample as possible. However, the present study shows that also a number of well isolated prompt electrons is removed by this cut. In Figure 7.18 (b) we observe that most of the electrons in this  $\Delta R$  region are actually

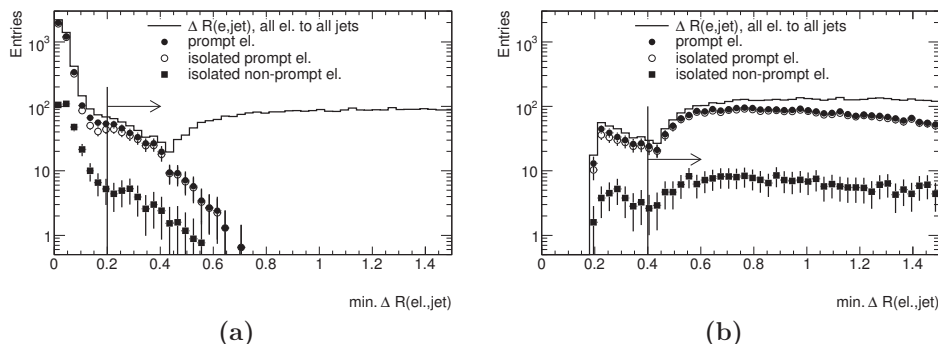


Figure 7.18:  $\Delta R$  between electrons and jet objects (a) without overlap removal (b) with overlap removal. The histograms show electrons in SU3, but the distributions for other SUSY samples and for  $t\bar{t}$  are similar. The histograms are normalized to  $1\text{fb}^{-1}$ .

*prompt* electrons and they do pass the isolation requirement. In SU3, using the test isolation  $E_{\text{cal}}^{\Delta R=0.2} < 6$  GeV, approximately 5% of all isolated prompt electrons are within  $0.2 < \Delta R < 0.4$ . At the same time, the fraction of all isolated *extra* electrons found in this region is 7% and 5% in SU3 and in  $t\bar{t}$ , respectively<sup>7</sup>. Applying the  $\Delta R(e, \text{jet}) > 0.4$  separation cut would reject both *prompt* and *extra* electrons from the events. It can be questioned whether the loss of prompt electrons is too high compared to the gain in terms of purity of the electron sample.

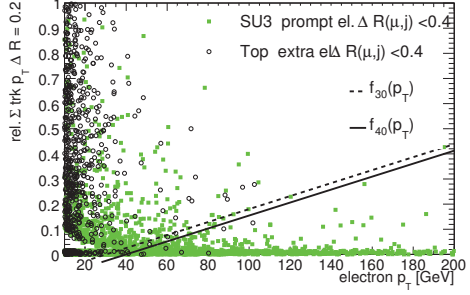
As an alternative to simply rejecting all electrons which fail the  $\Delta R(e, \text{jet}) > 0.4$  requirement, different cuts have been investigated. The aim was to reject *extra* electrons and still keep the prompt signal electrons. One possibility is to cut on the significance of the transverse impact parameter  $|d_0/\sigma(d_0)|$  where  $d_0$  is calculated with respect to the nominal primary interaction vertex which in MC is set to the coordinate system origin. Applying a standard cut  $|d_0/\sigma(d_0)| < 6$  on the electrons within  $0.2 < \Delta R < 0.4$ , one would keep about 90 % or more of the isolated prompt electrons from SUSY, while keeping around 70 % of the *extra* electrons from  $t\bar{t}$  (64 % requiring  $d_0/\sigma(d_0) < 5$ ). This cut has a stronger rejection power for *extra* electrons in  $t\bar{t}$  as these most often originate from semi-leptonic b-decays.

Another possibility is a combination of a higher  $p_T$  threshold with an isolation requirement through a  $p_T$ -dependent linear cut. This method has been developed in more detail in Section 7.2.3 devoted to muons, for which it is most relevant. However, it is interesting to note that the same cut is also very efficient for electrons. Figure 7.19 shows  $p_T$  along the x-axis and the relative  $p_{T\Sigma\text{track}}^{\Delta R=0.2}$  along the y-axis for prompt electrons in SU3 and *extra* electrons in  $t\bar{t}$  within a jet  $0.2 < \Delta R(e, \text{jet}) < 0.4$ . The lines  $f_{30}(p_T)$  and  $f_{40}(p_T)$ , given by Equations 7.1 and 7.2 in Section 7.2.3, indicate a cut which imposes a  $p_T$  threshold at 30 GeV and 40 GeV. Electrons below the line are accepted.

The study in this section suggests that the cuts developed for muons within a jet can also be applied for electrons. The effect of these on electron efficiency cuts will be further discussed in Section 7.5.

<sup>7</sup>The fractions are comparable in the other SUSY samples.

Figure 7.19: Scatter plot of the relative track-based isolation  $p_{T\Sigma\text{track}}^{\Delta R=0.2}$  along the y-axis vs.  $p_T$  for prompt electrons in SU3 and *extra* electrons in  $t\bar{t}$ .



## 7.4 Isolation optimization for tripleton signal

The main aim of this study is to focus on tripleton events and compare the signal selection efficiency and background rejection factor for different isolation criteria and possibly optimize the cuts. Two SM processes, the  $Zb$  and  $t\bar{t}$  production can have in final states up to two prompt leptons and  $b$ -jets and thus a high probability for creating signal-like events through the presence of *extra* leptons. These processes are particularly challenging because of the very high cross sections, 450 pb for  $t\bar{t}$  and 154 pb for  $Zb$ , as compared to 27.7 pb and 7.2 pb for SU3 and SU2, respectively. The isolation requirement is therefore an efficient background rejection tool.

Leptons are selected according to the object definition in Section 7.3 and 7.2. The events are required to have 3 or more leptons and the signal selection efficiency and background rejection factor are calculated for different isolation criteria. Tripleton events from SUSY are considered as signal, while tripleton events from  $t\bar{t}$ ,  $Zb$  and diboson events are counted as background. SU3 is chosen to represent SUSY in all figures in this section.

Signal efficiency  $\epsilon_S$  and background rejection factor  $r_B$  are defined as:

$$\epsilon_S = \frac{\text{\#Selected Signal 3-lep events}}{\text{\#Total Signal 3-lep events}} \quad (7.3)$$

$$r_B = 1 - \frac{\text{\#Selected Background 3-lep events}}{\text{\#Total Background 3-lep events}} \quad (7.4)$$

Since this analysis is aiming at optimizing the cuts for discovery, the statistical significance, given by

$$S_0 = \frac{S}{\sqrt{B}} \quad (7.5)$$

is the appropriate figure of merit.  $S$  is the number of signal events while  $B$  is the number of background events.  $S_0$  is plotted as function of the isolation cut and the value giving the maximum significance is selected. These plots give also a useful comparison between the different isolation variables. The  $p_T$ -dependent isolation cut  $f_{40}(p_T)$  given by Eq. 7.2, is applied on leptons inside a jet ( $\Delta R(l, jet) < 0.4$ ) throughout this section.

Figure 7.20 is a comparison of the significance  $S_0$  and  $\epsilon_S$  vs.  $r_B$  for the calorimeter based (a-b) and track-based isolation (c-d). Figure 7.21 shows the same comparison for the relative isolation. In both cases, the highest significance is obtained with the track-based isolation  $p_{T\Sigma\text{track}}^{\Delta R=0.3}$ .

The plots of  $\epsilon_S$  vs.  $r_B$  reveal a general characteristic of the different cone widths. For a given efficiency, typically above approximately 0.7, the  $\Delta R = 0.2$  cone provides the

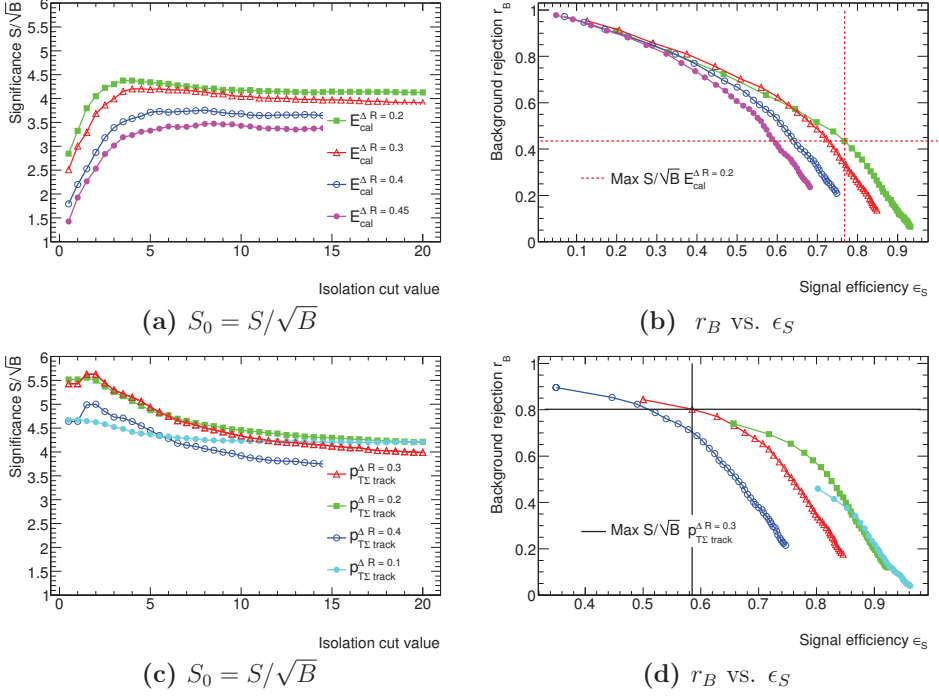


Figure 7.20: Significance  $S_0$  as a function of isolation cut value and  $\epsilon_S$  vs.  $r_B$  for trilepton events from SU3 and from SM background. Plots (a) and (b) show the results for the calorimeter-based isolation  $E_{\text{cal}}^{\Delta R=0.X}$  while plots (c) and (d) for the track-based isolation variables  $p_{T\Sigma\text{ track}}^{\Delta R=0.X}$ .

highest background rejection. The results for the track-based isolation energy calculated in the cone with  $\Delta R = 0.1$ ,  $p_{T\Sigma\text{ track}}^{\Delta R=0.1}$  are also presented in figure 7.20 and we see that the narrow cone can provide the highest signal efficiency  $\epsilon_S$ , but for the same loose cuts also the lowest background rejection. For lower efficiencies  $\epsilon_S$ , the wider cones give the highest background rejection. For the trilepton signature where both the signal and background events have a high jet activity, it looks like the track-based isolation calculated in the wider cone with  $\Delta R = 0.3$  provides high efficiency for signal and it is wide enough to include jet constituents in case of leptons originating from jets. This feature of the wider cones is mirrored in the significance scan (Figure 7.20 (c)) where the isolation in the  $\Delta R = 0.3$  cone gives the highest significance.

An isolation criterion which cuts on the  $p_T$  of the most energetic track in a cone,  $p_{T\text{max track}}^{\Delta R=0.X}$ , defined in Section 7.2.2, shows slightly lower significances than the  $p_{T\Sigma\text{ track}}^{\Delta R=0.X}$  isolation. A cut on the highest  $p_{T\text{max track}}^{\Delta R=0.X}$  of all leptons in the event, making it a global event isolation, was also tested. However, this method gives poorer results than when the cut on  $p_{T\text{max track}}^{\Delta R=0.X}$  is applied on the individual leptons.

A comparison of the maximum values of  $S_0$ , Figures 7.20 (a) and 7.21 (a), shows that the relative isolation variables give better results. For the calorimeter-based isolation, both the original  $E_T$  sum and the relative one, the  $\Delta R = 0.2$  cone gives the highest significances. For the track-based isolation,  $\Delta R = 0.3$  is preferable. The optimization is

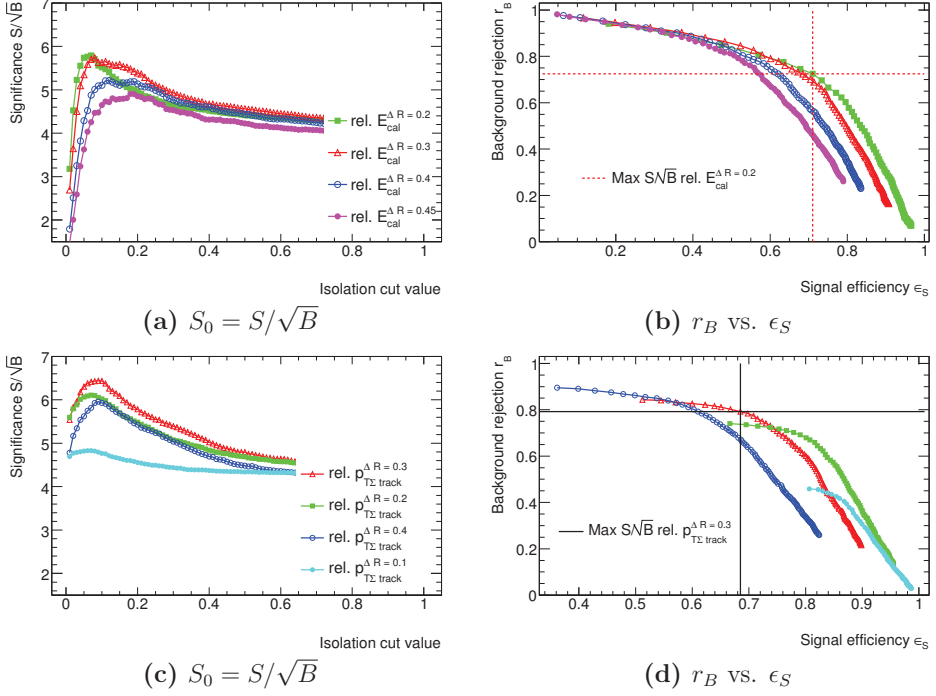


Figure 7.21: Same as Figure 7.20, but for relative calorimeter-based isolation  $E_{\text{cal}}^{\Delta R=0.X}$  (a),(b) and relative track-based isolation variables  $p_{T\Sigma \text{ track}}^{\Delta R=0.X}$  (c),(d).

similar for the other SUSY samples.

It must be mentioned that once the  $f_{40}(p_T)$  cut is applied on the leptons inside jets, most of the non-isolated muons are rejected as 90-95% of such muons are found close to a jet. The isolation cut is therefore most crucial for selecting isolated electrons.

The results of the isolation and lepton-jet separation discussion will now be combined in a study of single lepton efficiencies and fake rates.

## 7.5 Single lepton selection efficiency and fake rate

The observation of the trilepton signal requires a good single lepton selection efficiency as well as low fake rate. In particular, the dominant backgrounds  $t\bar{t}$  and  $Zb$  are the ones that are considerably reduced by a strong isolation criterion. When real ATLAS data are available, efficiencies and fake rates will have to be calculated using data driven methods. However, a Monte Carlo study may provide important qualitative results regarding the performance of the different isolation criteria. This study also investigated the impact of the lepton-jet separation requirement and the suggested  $p_T$ -dependent isolation cut for leptons in jets given by Eq. 7.2 in Section 7.2.3. The study shows that by applying this cut in the plane spanned by the relative  $p_{T\Sigma \text{ track}}^{\Delta R=0.2}$  and  $p_T$ , one can obtain a smaller efficiency loss compared to a simple rejection of all leptons close to jets, while at the same time keeping a low fake rate.

Thus three possible treatments of leptons in jets have been studied:

- **No cuts** - showed as a reference
- **Lepton-jet veto** -  $\Delta R(l, jet) > 0.4$  separation requirement. Leptons closer to a jet than  $\Delta R(l, jet) = 0.4$  are discarded (CSC convention)
- **Linear isolation cut** - leptons within a jet are subject to linear  $p_T$ -isolation cut given by the line:  $f_{40}(p_T) = 0.003(p_T - 40 \text{ GeV})$ , introduced in Eq. 7.2

No other isolation requirements were imposed at this point as the aim was to study the effect of the cuts on leptons in jets only.

The lepton reconstruction efficiency is defined as

$$\mathcal{E}_\ell \equiv \frac{n_\ell^{\text{prompt MC matched to Reco}}}{n_\ell^{\text{prompt MC}}}, \quad (7.6)$$

where the numerator is the number of *generator-level leptons* matched to a reconstructed candidate.  $n_\ell^{\text{MC}}$  is the total number of generator-level leptons which includes direct decay products of SUSY particles (gauginos and sleptons) and decays of SM gauge bosons and taus. The efficiency is calculated for leptons with minimum  $p_T$  of 10 GeV and to avoid the effect of limited  $p_T$  resolution, a lower threshold of 5 GeV was set for the reconstructed candidates. No isolation cut was applied on the generator-level leptons. A match was accepted if a reconstructed candidate was found within  $\Delta R < 0.02$ .

|                | No cuts        | $\Delta R(l, jet) > 0.4$ | $f_{40}(p_T)$ cut if $\Delta R(l, jet) < 0.4$ |
|----------------|----------------|--------------------------|---|
| SU3 $e$        | 71.0 $\pm$ 0.4 | 66.9 $\pm$ 0.5           | 69.2 $\pm$ 0.2                                |
| $t\bar{t} e$   | 70.9 $\pm$ 0.3 | 68.7 $\pm$ 0.3           | 69.2 $\pm$ 0.2                                |
| SU3 $\mu$      | 91.8 $\pm$ 0.3 | 70.4 $\pm$ 0.4           | 80.9 $\pm$ 0.2                                |
| $t\bar{t} \mu$ | 92.4 $\pm$ 0.2 | 72.4 $\pm$ 0.3           | 82.4 $\pm$ 0.1                                |

Table 7.3: Electron and muon efficiency [%] for different treatments of leptons inside jets. *No cuts* refers to option implying no requirements on leptons in jet. No other isolation cuts were applied.

Due to the overlap removal, the lepton-jet separation cut has only a small effect on the electrons. Nevertheless, as shown in Table 7.3 and Figures 7.22 (a) and (b), introducing the  $p_T$ -dependent isolation cut, given by  $f_{40}(p_T)$ , instead of rejecting the electrons with  $0.2 < \Delta R(e, jet) < 0.4$  gives slightly higher efficiency. In the case of muons, the selection efficiency is significantly reduced by the lepton-jet veto, for most samples by as much as 20%. By applying the  $f_{40}(p_T)$ -isolation cut, one can recover about 10% of the efficiency. Figure 7.22 (c) shows that the improvement is specially notable for muons with high  $p_T$ . The  $\eta$ -distribution, Figure 7.22 (d), shows that the improvement is most significant in the central barrel region. This is in good agreement with the study of the  $\Delta R_{min}(\mu, jet)$  shown in Figure 7.9.

The fake rate is defined as

$$\mathcal{F}_\ell \equiv \frac{n_\ell^{\text{Reco not-matched to MC}}}{n_{\text{jet}}^{\text{MC}}}, \quad (7.7)$$

where the numerator is the number of reconstructed leptons which are *not* matched to a generator-level lepton and  $n_{\text{jet}}^{\text{MC}}$  is the number of generator-level jets. These jets are

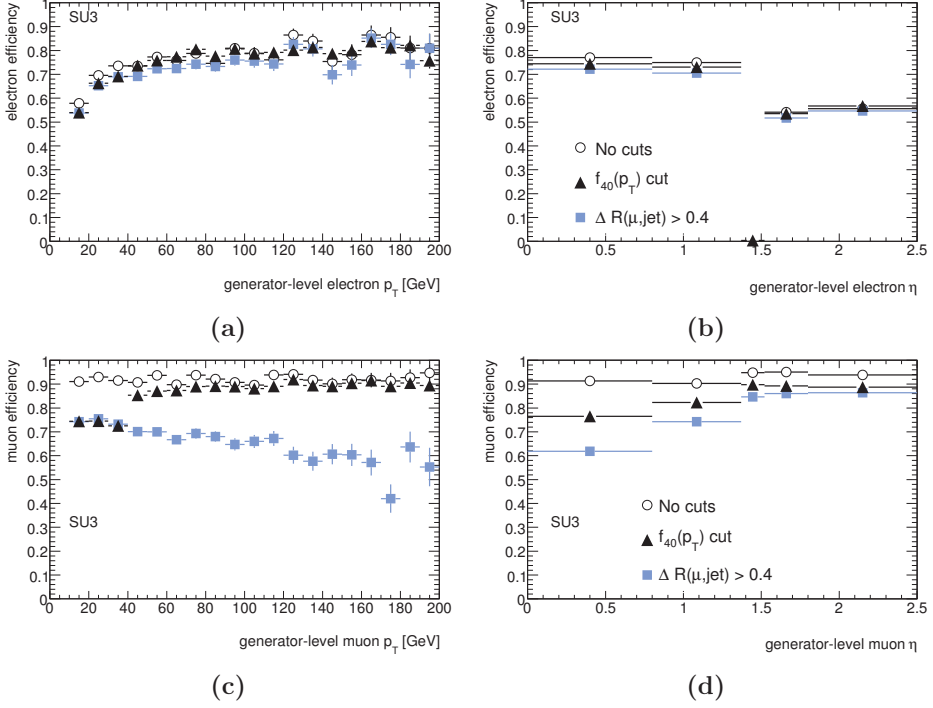


Figure 7.22: The efficiency for electrons (a,b) and muons (c,d) applying different treatment of leptons within a jet. All distributions are for the SU3 sample. The plots are shown as a function of the  $p_T$  (a,c) and  $|\eta|$  (b,d) of the matched Monte Carlo lepton.

required to be in the central barrel region,  $|\eta| < 2.5$ , and have energy <sup>8</sup>  $E > 7$  GeV. If a jet overlaps with an electron within  $\Delta R < 0.2$  it is rejected. As no other overlap rejection is applied, the jets include photons, hadronic tau-jets and b-jets. The fake rate is calculated for leptons with  $p_T > 10$  GeV and in order not to miss a match a lower threshold of  $p_T > 5$  GeV is applied for the generator-level leptons.

|                | No cuts        | $\Delta R(l, jet) > 0.4$ | $f_{40}(p_T)$ cut if $\Delta R(l, jet) < 0.4$ |
|----------------|----------------|--------------------------|---|
| SU3 $e$        | $3.1 \pm 0.1$  | $2.8 \pm 0.1$            | $3.0 \pm 0.1$                                 |
| $t\bar{t} e$   | $4.3 \pm 0.1$  | $4.0 \pm 0.1$            | $4.1 \pm 0.1$                                 |
| SU3 $\mu$      | $12.0 \pm 0.2$ | $0.49 \pm 0.04$          | $0.51 \pm 0.04$                               |
| $t\bar{t} \mu$ | $18.2 \pm 0.2$ | $1.1 \pm 0.1$            | $1.1 \pm 0.1$                                 |

Table 7.4: Electron and muon fake rate  $[\times 10^{-3}]$  for different treatments of leptons inside jets. *No cuts* refers to option implying no requirements on leptons in jet. No other isolation cuts were applied.

The fake rate was studied for the three different treatments of leptons in jets and the results in Table 7.4 show that the lepton-jet veto (column 3) dramatically reduces the

<sup>8</sup> $E > 7$  GeV is the minimum imposed by the reconstruction algorithm.



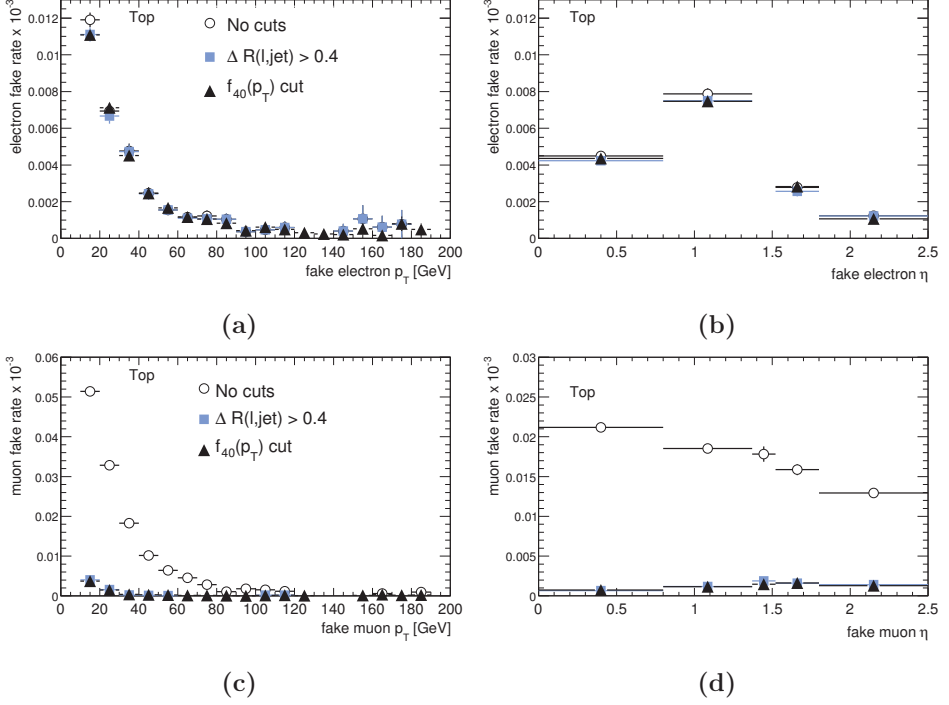


Figure 7.23: The fake rate for electrons (a,b) and muons (c,d) applying different treatment of leptons within a jet, for the  $t\bar{t}$  sample. The plots are shown as a function of the  $p_T$  (a,c) and  $\eta$  (b,d) of the fake lepton.

muon fake rate, as expected, since 90-95% of all *extra* muons are found close to a jet. Replacing the naive rejection of such muons with the  $f_{40}(p_T)$ -isolation cut leads to only a small increase of the fake rate (Table 7.4, column 4). These cuts have only a minor effect on the electron fake rate as they concern only a small fraction of the electron sample. Nevertheless, the fake rate is not affected too much when replacing the lepton-jet veto with the  $p_T$ -dependent isolation cut. The electron and muon fake rate in  $t\bar{t}$  is shown in Figure 7.23 as function of  $p_T$  and  $\eta$ .

This study shows the good performance of the  $p_T$ -dependent isolation cut. Its power to reject the *extra* leptons is comparable to the veto of leptons inside jets, while at the same time, the selection efficiency is kept higher than for the veto.

### 7.5.1 Performance of the different isolation requirements

The  $p_T$ -dependent isolation cut for leptons inside jets  $f_{40}(p_T)$  (see Eq. 7.2) has been chosen as the baseline for further study of two isolation criteria which are to be applied on all leptons

- Calorimeter-based relative isolation:  $E_{\text{cal}}^{\Delta R=0.2}/p_T < 0.08$
- Track-based relative isolation:  $p_{T\Sigma\text{track}}^{\Delta R=0.3}/p_T < 0.11$

|  | SU1 $e$        | SU2 $e$        | SU3 $e$        | SU4 $e$        | $t\bar{t}e$    |
|--|----------------|----------------|----------------|----------------|----------------|
| $f_{40}(p_T)$ -cut only                              | 69.8 $\pm$ 0.2 | 66.8 $\pm$ 0.5 | 69.2 $\pm$ 0.2 | 67.2 $\pm$ 0.4 | 69.2 $\pm$ 0.1 |
| Rel. $E_{\text{cal}}^{\Delta R=0.2} < 0.08$          | 64.0 $\pm$ 0.2 | 60.4 $\pm$ 0.5 | 63.3 $\pm$ 0.2 | 59.4 $\pm$ 0.3 | 64.1 $\pm$ 0.2 |
| Rel. $p_{T\Sigma\text{track}}^{\Delta R=0.3} < 0.11$ | 64.2 $\pm$ 0.2 | 60.8 $\pm$ 0.5 | 62.8 $\pm$ 0.2 | 59.2 $\pm$ 0.2 | 64.1 $\pm$ 0.2 |
|  | SU1 $\mu$      | SU2 $\mu$      | SU3 $\mu$      | SU4 $\mu$      | $t\bar{t}\mu$  |
| $f_{40}(p_T)$ -cut only                              | 81.6 $\pm$ 0.2 | 82.0 $\pm$ 0.4 | 80.9 $\pm$ 0.2 | 78.8 $\pm$ 0.3 | 82.4 $\pm$ 0.1 |
| Rel. $E_{\text{cal}}^{\Delta R=0.2} < 0.08$          | 78.4 $\pm$ 0.2 | 78.8 $\pm$ 0.4 | 77.4 $\pm$ 0.2 | 75.3 $\pm$ 0.2 | 79.0 $\pm$ 0.1 |
| Rel. $p_{T\Sigma\text{track}}^{\Delta R=0.3} < 0.11$ | 77.8 $\pm$ 0.2 | 78.3 $\pm$ 0.4 | 76.2 $\pm$ 0.2 | 74.0 $\pm$ 0.2 | 80.0 $\pm$ 0.1 |

Table 7.5: Electron and muon efficiency [%] for different isolation criteria.

|  | SU1 $e$         | SU2 $e$         | SU3 $e$         | SU4 $e$         | $t\bar{t}e$     |
|--|-----------------|-----------------|-----------------|-----------------|-----------------|
| $f_{40}(p_T)$ -cut only                              | 2.2 $\pm$ 0.2   | 2.2 $\pm$ 0.1   | 3.0 $\pm$ 0.1   | 3.6 $\pm$ 0.1   | 4.1 $\pm$ 0.1   |
| Rel. $E_{\text{cal}}^{\Delta R=0.2} < 0.08$          | 0.36 $\pm$ 0.02 | 0.37 $\pm$ 0.04 | 0.70 $\pm$ 0.02 | 0.39 $\pm$ 0.02 | 0.61 $\pm$ 0.02 |
| Rel. $p_{T\Sigma\text{track}}^{\Delta R=0.3} < 0.11$ | 0.53 $\pm$ 0.02 | 0.38 $\pm$ 0.03 | 0.99 $\pm$ 0.03 | 0.42 $\pm$ 0.02 | 0.62 $\pm$ 0.02 |
|  | SU1 $\mu$       | SU2 $\mu$       | SU3 $\mu$       | SU4 $\mu$       | $t\bar{t}\mu$   |
| $f_{40}(p_T)$ -cut only                              | 0.65 $\pm$ 0.02 | 0.72 $\pm$ 0.05 | 0.51 $\pm$ 0.04 | 1.1 $\pm$ 0.1   | 1.1 $\pm$ 0.1   |
| Rel. $E_{\text{cal}}^{\Delta R=0.2} < 0.08$          | 0.15 $\pm$ 0.01 | 0.29 $\pm$ 0.03 | 0.15 $\pm$ 0.01 | 0.38 $\pm$ 0.02 | 0.43 $\pm$ 0.02 |
| Rel. $p_{T\Sigma\text{track}}^{\Delta R=0.3} < 0.11$ | 0.13 $\pm$ 0.01 | 0.20 $\pm$ 0.02 | 0.13 $\pm$ 0.01 | 0.19 $\pm$ 0.01 | 0.22 $\pm$ 0.01 |

 Table 7.6: Electron and muon fake rate [ $\times 10^{-3}$ ] for different isolation criteria.

The choice of the isolation variable is based on the study in Section 7.4 which showed that the cone of  $\Delta R = 0.2$  gave best results for the calorimeter-based isolation while  $\Delta R = 0.3$  was best for the track-based. The cut value is a result of the optimization of the signal significance for the trilepton analysis presented in Section 7.4.

The results in Table 7.5 show interesting features of these two isolation criteria. Requiring them reduces the selection efficiency by approximately 6% for electrons and 4% for muons, compared to the results when only applying the  $p_T$ -dependent  $f_{40}(p_T)$ -cut on leptons inside jets, (referred to as  $f_{40}(p_T)$ -cut). The efficiency for calorimeter- and track-based isolation is comparable for electrons while for muons, the calorimeter-based isolation gives slightly higher efficiency, except for the  $t\bar{t}$  sample.

As shown in Table 7.6, the isolation cuts reduce the fake rate by a factor 3-6 and the impact is largest on the electrons since most of the *extra* muons are already removed by the  $f_{40}(p_T)$ -cut. The lowest fake rate is obtained for electrons with the calorimeter-based isolation while for muons with the track-based.

Figure 7.24 (a) shows that isolation cuts reduce significantly the electron fake rate in the low  $p_T$  region. The  $\eta$ -bin  $[0.8, 1.37]$  corresponds to the transition region between the barrel and the extended barrel hadronic calorimeter and the reduced performance of the hadronic leakage cut is mirrored in the increased electron fake rate.

The muon fake rate is already strongly suppressed by the  $f_{40}(p_T)$ -cut. However it can be noted that the track-based isolation performs better in the transition region between the barrel and the forward parts of the Inner Detector  $\eta = [1.37, 1.8]$ , see Figure 7.24 (d).

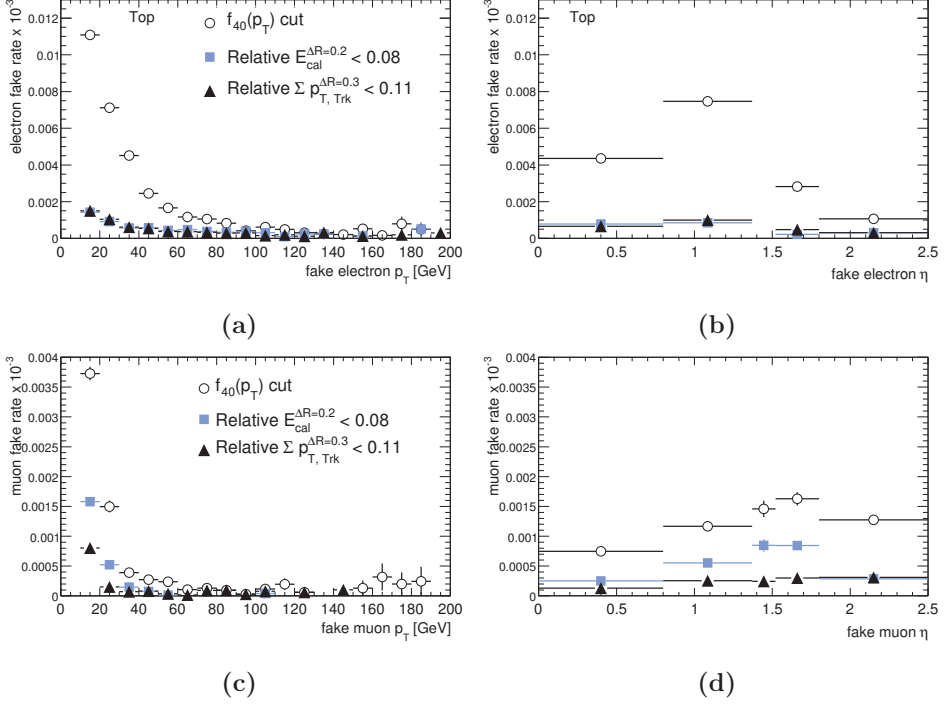


Figure 7.24: The fake rate for electrons (a,b) and muons (c,d) applying different isolation criteria, for the  $t\bar{t}$  sample. The plots are shown as a function of the  $p_T$  and  $\eta$  of the the fake electron.

The purity of the lepton samples has also been studied for the different isolation criteria. Purity is defined as

$$\mathcal{P}_\ell = \frac{n_\ell^{\text{Reco match to MC}}}{n_\ell}, \quad (7.8)$$

where the numerator is the number of reconstructed leptons matched to a generator-level lepton, while  $n_\ell$  is the total number of selected reconstructed leptons. The effect of the  $p_T$ -dependent isolation cut  $f_{40}(p_T)$  on the electron purity is small as the cut concerns only the small fraction of electrons. After this cut, purity is measured to be approximately 87% and 90% for SU3 and  $t\bar{t}$ , respectively. The muon purity, on the other hand, improves significantly with the  $f_{40}(p_T)$  cut. It increases by 28% and 23% for SU3 and  $t\bar{t}$ , respectively, and before any other isolation cuts, it is already at the level of about 98% for both SU3 and  $t\bar{t}$ .

When in addition the cut on relative  $p_{T\Sigma}^{\text{AR}=0.3} < 0.11$  is applied on all leptons, not only those close to a jet, a purity of 94% and 98% is also obtained for electrons in SU3 and  $t\bar{t}$ , respectively. The improvement for muons is much smaller, only 1%. The purities for the other SUSY samples are comparable, differing by not more than around 1%.

These quoted results so far are based on an analysis of events with no requirement on the number of leptons in the event. It is interesting to note that as soon as three or more leptons are required in an event the purity changes significantly. For the SUSY signal

samples the lepton purity is unchanged or even better. Using SU3 as example, the purity is now 96% and 99% for electrons and muons respectively. While for  $t\bar{t}$  events, which by definition can have at most two prompt leptons, the only way of passing the trilepton requirement is through the presence of a non-prompt lepton. After imposing the trilepton requirement, the electron purity is degraded to 70% and to 78% for muons.

# Chapter 8

## SUSY searches, trilepton analysis at 14 TeV

Trilepton events are defined as final states with three or more isolated leptons. Such SUSY final states can be divided into two categories: jet inclusive, where the leptons are accompanied by a number of high  $p_T$  jets and jet exclusive, where the leptons come from gauginos produced in the hard process. In the first category leptons occur as decay products in a long decay chain starting with a squark or gluino. This kind of processes have high cross section in benchmark points where squarks and gluinos are relatively light like in SU3 and SU4. A Feynman diagram for such a process was shown in Figure 6.12 (b). At the LHC, pure QCD-processes will have very high cross sections, and requiring leptons in an event, especially as many as three, strongly suppresses QCD and other SM backgrounds.

The so called “Focus point region” in the mSUGRA parameter space (exemplified with SU2) differs from other points by having very high squark and slepton masses, while the gluinos<sup>1</sup>, and more important from the point of view of this analysis, the gauginos are sufficiently light. In this point the direct production of chargino and neutralino pairs, shown in Figure 6.11 (b), accounts for the largest fraction of the total cross section. In case of a very heavy scalar scenario this could even be a discovery channel.

Searches for SUSY in trilepton final states follow therefore these two strategies. In both cases, high selection efficiency of well isolated leptons combined with a good rejection of *extra* leptons is crucial.

### 8.1 Object definition

The first step of the analysis is to require at least three well isolated leptons. Leptons are selected according to the object definitions in Sections 7.2 (muons) and 7.3 (electrons). Taking into account the results of the study of the lepton-jet separation, a lepton which is found closer to a jet axis than  $\Delta R = 0.4$ , is subject to the  $p_T$ -dependent isolation cut  $f_{40}(p_T)$  given in Eq. 7.2. Below follows a short summary of the object selection:

---

<sup>1</sup>However, in some parameter space regions, gluinos may also be very heavy and possibly out of the LHC energy reach.

- **Electrons:**

- $p_T > 10$  GeV and  $|\eta| < 2.5$ , events containing an electron with  $|\eta| \in [1.37, 1.52]$ , corresponding to the transition region between the barrel and the endcap, are rejected
- “medium electron”, reconstructed by the **egamma** algorithm described in Section 7.3
- Relative track-based isolation:  $p_{T\Sigma\text{track}}^{\Delta R=0.2}/p_T < 0.11$  and if close to a jet ( $\Delta R(e, jet) < 0.4$ ) linear cut  $f_{40}(p_T)$  given by Eq. 7.2 in Section 7.2.3

- **Muons:**

- $p_T > 10$  GeV and  $|\eta| < 2.5$
- Staco combined muons and best match requirement,  $\chi^2 < 100$ , **isHighPt** - high  $p_T$ -muons as described in Section 7.2
- Relative track-based isolation:  $p_{T\Sigma\text{track}}^{\Delta R=0.2}/p_T < 0.11$  and if close to a jet ( $\Delta R(e, jet) < 0.4$ ) linear cut  $f_{40}(p_T)$  given by Eq. 7.2 in Section 7.2.3

- **Jets:**

- $p_T > 10$  GeV and  $|\eta| < 2.5$
- **Cone4Tower** seeded cone algorithm with  $\Delta R = 0.4$  described in Section 7.1
- Jet is removed if overlapping with an electron within  $\Delta R < 0.2$

The  $\cancel{E}_T$  is the refined summation **MET\_RefFinal**.

In addition, if the invariant mass of a combination of same flavor opposite sign ( $l^+l^-$ ) leptons was less than 20 GeV, the two leptons were rejected. This was done in order to remove lepton pairs from photon conversions and light resonances, such as  $J/\psi$ ,  $\Upsilon$ , etc.

The final choice of the isolation cut  $p_{T\Sigma\text{track}}^{\Delta R=0.3}/p_T < 0.11$  is based on a comparison of the statistical significance  $S/\sqrt{B}$  for this track-based isolation and the calorimeter-based alternative  $E_{\text{cal}}^{\Delta R=0.2}/p_T$  presented in Section 7.4. The cut value gives optimal or close to optimal results for all SUSY samples.

Figure 8.1 (a) shows the lepton multiplicity applying the definitions described above. In the 3- and 4-lepton bin the dark area representing the diboson background <sup>2</sup> is visible above the grey  $Zb + t\bar{t}$ -area indicating that diboson backgrounds become more important. Figure 8.1 (b) shows the jet multiplicity in trilepton events. SUSY, with the exception of the direct gaugino production labelled  $\text{SU2} - \chi\chi$ , is clearly distinguishable by its high jet multiplicity. It should be noted that the minimum jet- $p_T$  cut of 10 GeV is low compared to most common SUSY event selections.

The trilepton event selection is based on a study of three key distributions: lepton and jet  $p_T$  and missing energy,  $\cancel{E}_T$ . From Figure 8.2 which shows the lepton  $p_T$  distributions in events with  $N_l \geq 3$ , we see that the SUSY-leptons have a relatively hard  $p_T$  spectrum. Looking at the  $p_T$  distribution for the third leading lepton we see that the SUSY-leptons are more energetic than those from  $t\bar{t}$  and  $Zb$  which are often *extra* leptons, typically soft

<sup>2</sup>The plots in this chapter show the backgrounds as a stack plot with the SUSY-signal distributions superimposed.

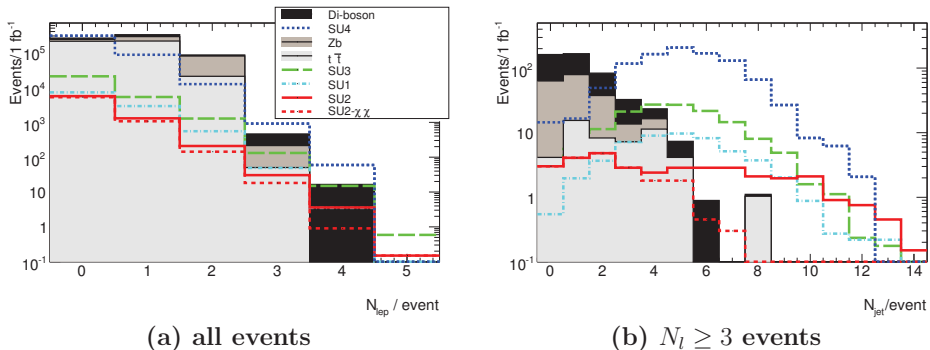


Figure 8.1: Lepton multiplicity (a) and jet multiplicity after requiring 3 or more isolated leptons (b).

leptons from b-decays. The high- $p_T$  region of the background distribution in Figure 8.2 (c) is dominated by diboson events.

### 8.1.1 Significance measures

The SUSY searches presented in this and the next chapters are typical examples of a cut and count analysis where the main question is whether it is possible to claim an excess of events over the expected SM background. Intuitively one would like to have a measure which can discriminate between a statistical fluctuation of the background and a possible presence of a real signal. In the following, the number of signal events is denoted by  $S$ , the number of background events by  $B$ . One of the most commonly used significance measures:

$$S_0 = \frac{S}{\sqrt{B}} \quad (8.1)$$

is selected as the figure of merit for the cut optimization and will represent the main results. Although it overestimates the discovery significance and does not take into account the uncertainty of the background, it is simple and useful for cut optimization [165]. It gives also a rough estimate of the discovery potential. However, its validity is limited to the Gaussian cases where the number of signal and background events is significantly larger than unity. This is unfortunately not the case for all of the analyses presented here. In a Poissonian regime it is only an approximation in cases where  $S \ll B$ .

The second significance measure which will be quoted is given by the simple expression

$$S_N = \frac{S}{\sqrt{B + (\sigma_B)^2}} \quad (8.2)$$

where  $\sigma_B$  represents the background uncertainty. The uncertainties on the backgrounds relevant to the present analysis have been estimated to be 20% by the SM-working group [78].

Another estimate which also takes the background uncertainty into account is the  $Z_N$  significance and it is valid in situations when the number of events is not necessarily Gaussian. The probability  $p$  that the background fluctuates to the observed signal is derived from convolution of the Poisson probability density function (pdf) which account for the

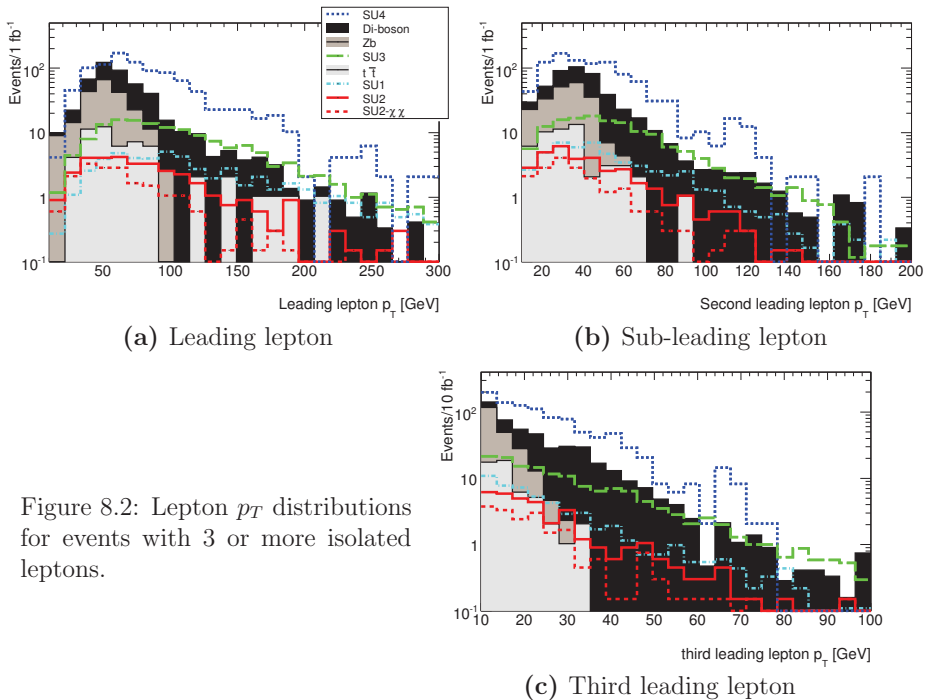


Figure 8.2: Lepton  $p_T$  distributions for events with 3 or more isolated leptons.

statistical error with a Gaussian pdf to include the effect of non negligible systematic uncertainties.  $p$  is then transformed into “standard deviations” using the formula:

$$Z_N = \sqrt{2} \operatorname{erf}^{-1}(1 - 2p) \quad (8.3)$$

A detailed description of  $Z_N$  can be found in [78], page 1590-1591. The calculation of  $Z_N$  uses the ROOT [90] package StatTools<sup>3</sup>.

## 8.2 Jet inclusive analysis

The event selection in this analysis is based on a set of simple and robust cuts which can be applicable already during the early data taking. The trilepton requirement strongly reduces the SM background and most of the accepted SUSY events are characterised by at least one high- $p_T$  jet and large  $\cancel{E}_T$ . These two features clearly distinguish the SUSY signal from the SM-background as can be seen in Figure 8.3 (a) showing the leading jet  $p_T$  and Figure 8.5 (a) showing  $\cancel{E}_T$  for 3-lepton events.

The jet inclusive analysis follows two strategies: “3-leptons+jet” which uses jets in the event selection and “3-leptons+ $\cancel{E}_T$ ” which does not impose any criteria on the jets. During the early data taking there will be large uncertainties related to both jets and  $\cancel{E}_T$  so it can be an advantage to have two different methods.

<sup>3</sup>At the time when this analysis was done I did not have a stable version of the program for calculation of  $Z_N$ . The calculation is done later using the numbers in the tables and is therefore quoted with lower precision.



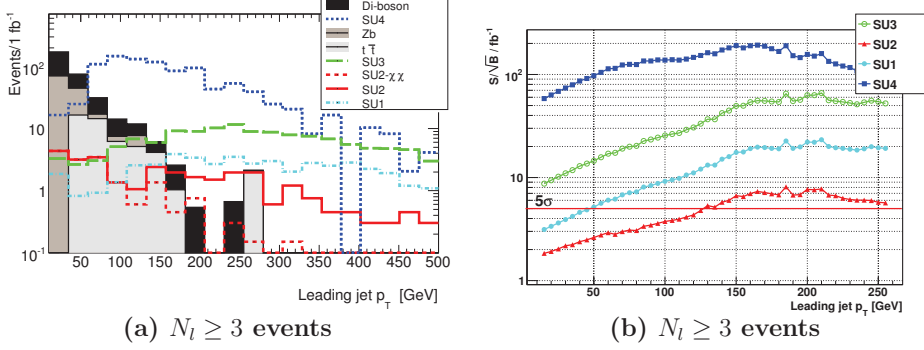


Figure 8.3: Leading jet  $p_T$  (a) after requiring 3 isolated leptons. Significance  $S/\sqrt{B}$  (b) for different minimum  $p_T$  cuts on the leading jet.

### 8.2.1 “3-leptons+jet”

Figure 8.3 (b) shows the  $S/\sqrt{B}$  as a function of the minimum leading jet  $p_T$  cut. For the SU4 sample the maximum value of  $S/\sqrt{B}$  is obtained for values around 150 GeV which is lower than for the other SUSY samples. This can be explained by SU4’s softer jet  $p_T$  spectrum, shown in Figure 8.3 (a), as compared to the other SUx samples. A cut of 200 GeV is selected as a baseline cut, both because it is a compromise between SU4 and the other samples for which higher jet  $p_T$  cuts give better results, and to allow a comparison with the results of the CSC trilepton inclusive analysis [78]. An alternative looser cut  $p_T^{jet1} > 150$  GeV will also be studied. It is favourable for points like SU4 while the significance for the other points is not very much lower as these seems to reach a plateau starting from around  $p_T^{jet1} = 150$  GeV. This looser cut allows also to keep a larger number of events such that the results are less sensitive to fluctuations caused by limited MC statistics and uncertainties related to the high values of jet  $p_T$ . As can be seen from Figure 8.4 (a) showing  $\cancel{E}_T$  after the cut  $p_T^{jet1} > 150$  GeV, the background is strongly reduced and the statistical significance does not improve with an additional cut on  $\cancel{E}_T$ .

Let us summarize the cuts:

- $N_l \geq 3$  : at least 3 isolated leptons
- $p_T^{jet1}$  -cut, two cut values are studied:
  - leading jet  $p_T^{jet1} > 200$  GeV
  - leading jet  $p_T^{jet1} > 150$  GeV

The effective mass  $M_{eff}$  is a useful quantity in SUSY searches as it is related to the SUSY mass scale. This analysis uses the following definition:

$$M_{eff} = \cancel{E}_T + \sum_{i=1}^{N_{jet}} p_T^{jet} + \sum_{i=1}^{N_{lep}} p_T^{lep} \quad (8.4)$$

where the  $p_T$  sum runs over all jets with  $p_T > 40$  GeV and all selected isolated leptons ( $p_T > 10$  GeV). Figure 8.4 (b) shows the  $M_{eff}$  distribution for trilepton events

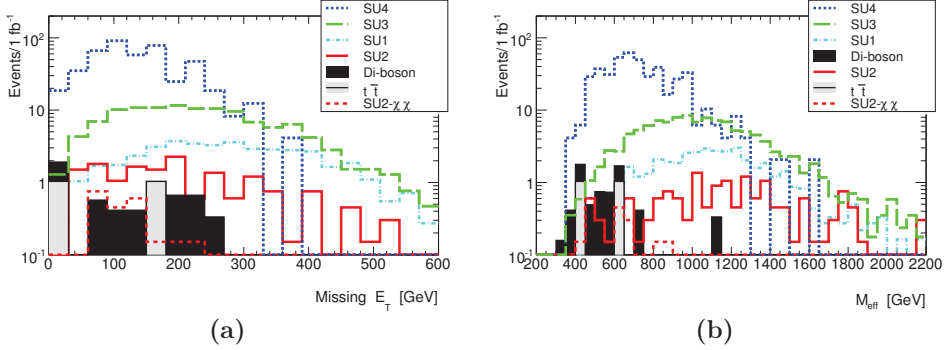


Figure 8.4: The  $\cancel{E}_T$  (a) and  $M_{eff}$  (b) distributions after requiring three or more isolated leptons and  $p_T^{jet1} > 150$  GeV.

after requiring  $p_T^{jet1} > 150$  GeV. One can see that the signal is clearly visible above the background.

Table 8.1 shows the cut flow for this analysis while Table 8.2 and 8.3 show the resulting statistical significances.

A significance of at least 5 is required in order to claim a discovery and the results in Tables 8.2 and 8.3 show that it would be feasible with less than  $1 \text{ fb}^{-1}$  of collected data for all studied SUSY points, possibly with the exception of SU2 where the main process is direct gaugino production for which a hard jet cut is not optimal. However, in this point there is also a considerable cross section for gluino production and the simple approach of the current analysis also applies.

Comparing the results of the tight cut  $p_T^{jet1} > 200$  GeV (Table 8.2) with the CSC trilepton inclusive searches [78] which only differs by the choice of the isolation requirement and the treatment of leptons close to a jet ( $\Delta R(l, jet) < 0.4$ ), one can see a clear improvement. The CSC publication quotes  $S/\sqrt{B}$  of 3.7, 27.1 and 90.0 for SU2, SU3 and SU4, respectively, which is more than 45% lower than the results of the present analysis (Table 8.2).

Applying the looser cut  $p_T^{jet1} > 150$  GeV (Table 8.3) one can still obtain promising results for all points, and for SU4 even improve. SU4 is the low mass point and the relatively smaller mass differences lead to softer  $p_T$  and  $\cancel{E}_T$  spectra.

The analysis has also been done with a looser requirement on leptons within the jet cone using the  $f_{30}(p_T)$  cut given by Eq. 7.1 which implies a lower minimum  $p_T$  cut at 30 GeV instead of  $f_{40}(p_T)$ . With the exception of SU2, all studied points have a significance well above the  $5\sigma$  limit, both for the  $p_T^{jet1} > 200$  GeV and  $p_T^{jet1} > 150$  GeV cut.

Knowing that there will be large uncertainties during the early data taking, the treatment of the leptons in the jet cone was simplified to only a tight  $p_T^{lep} > 40$  GeV and  $p_T^{lep} > 30$  GeV. One should keep in mind that all leptons are already subject to an isolation cut:  $p_{T\Sigma track}^{\Delta R=0.3}/p_T < 0.11$ . The resulting significances were reduced by 8-10 % but still showed promising discovery potential.

| Cut                           | SU1   | SU2  | SU3   | SU4    | $t\bar{t}$ | $Zb$   | WZ    | ZZ   | WW    | $Z\gamma$ | B      |
|-------------------------------|-------|------|-------|--------|------------|--------|-------|------|-------|-----------|--------|
| No Cuts                       | 10765 | 7150 | 27502 | 400777 | 444861     | 159395 | 15668 | 3859 | 40174 | 3287      | 667244 |
| $N_l \geq 3$                  | 53    | 34   | 150   | 982    | 52         | 165    | 199   | 57   | 2     | 5         | 480    |
| $p_T^{jet} > 200 \text{ GeV}$ | 36    | 12   | 102   | 252    | 1          | 0      | 1.3   | 0    | 0     | 0.2       | 2.6    |

Table 8.1: Cut-flow table. Number of events corresponds to the number after removal of events with electrons in crack. All numbers are normalized to  $1fb^{-1}$ .

Results “3-leptons+jet” inclusive analysis using the  $f_{40}(p_T)$ -cut given by Eq. 7.2. N/A means that  $Z_N$  is not calculable.

| Sample                      | SU1                              | SU2 | SU3  | SU4   | Sample                      | SU1                          | SU2 | SU3  | SU4   |
|-----------------------------|----------------------------------|-----|------|-------|-----------------------------|------------------------------|-----|------|-------|
| S                           | 36                               | 12  | 102  | 252   | S                           | 43                           | 16  | 120  | 463   |
| B                           | 2.6 (1 $t\bar{t}$ , 1.6 Diboson) |     |      |       | B                           | 6 (1 $t\bar{t}$ , 5 Diboson) |     |      |       |
| $S/B$                       | 13.7                             | 4.7 | 39.1 | 96.8  | $S/B$                       | 7.2                          | 2.7 | 20.4 | 78.5  |
| $S/\sqrt{B}$                | 22.1                             | 7.7 | 63.0 | 156.2 | $S/\sqrt{B}$                | 17.5                         | 6.6 | 49.4 | 190.7 |
| $S/\sqrt{B + (\sigma_B)^2}$ | 21.0                             | 7.3 | 60.0 | 148.7 | $S/\sqrt{B + (\sigma_B)^2}$ | 15.8                         | 6.0 | 44.4 | 171.5 |
| $Z_N$                       | 10                               | 5   | 21   | 38    | $Z_N$                       | 9                            | 4   | 19   | N/A   |

Table 8.2: Tight jet  $p_T$  cut  $p_T^{jet} > 200 \text{ GeV}$

Table 8.3: Loose jet  $p_T$  cut:  $p_T^{jet} > 150 \text{ GeV}$

### 8.2.2 “3-leptons+ $\cancel{E}_T$ ”

The cuts on  $\cancel{E}_T$  chosen for the study in the “3-leptons+ $\cancel{E}_T$ ”-analysis were set on the basis of the  $\cancel{E}_T$  distribution and corresponding significance plot of Figure 8.5 (b). Both  $\cancel{E}_T > 150 \text{ GeV}$  and  $\cancel{E}_T > 200 \text{ GeV}$  could be suitable for all points, except SU4 where the optimization favour  $\cancel{E}_T > 100 \text{ GeV}$ . After the cut at  $\cancel{E}_T > 150 \text{ GeV}$  the background from  $t\bar{t}$  and diboson are equal, while after the tighter cut at  $\cancel{E}_T > 200 \text{ GeV}$ , the latter is largest.

In order to reduce the backgrounds which involve  $Z$  bosons, all possible combinations of same flavor opposite sign lepton pairs,  $l^+l^-$ , were required to have an invariant mass outside the  $Z$ -mass window defined as  $|M_{l^+l^-} - M_Z| > 6 \text{ GeV}$ . A cut optimization study of the  $S/\sqrt{B}$  as function of the cut on the difference  $|M_{l^+l^-} - M_Z|$  shows that a cut at 6 GeV is favourable for all points. In Figure 8.6 (a) which shows the distribution of  $M_{l^+l^-}$  after requiring 3 isolated leptons, we can clearly see the narrow  $Z$  peak. For trilepton events, the  $Z$ -mass window cut removes 75% of the  $WZ$  and  $Zb$  background and 83% of the  $ZZ$  background, while the SUSY signals are reduced by about 28%. It is specially the SU3 point that is sensitive to changes in the size of the  $Z$ -mass window as the edge of the  $M_{l^+l^-}$  distribution is expected at 100.2 GeV. The cut at 6 GeV is looser than the often used mass window of 10 GeV, however combined with a tight requirement on  $\cancel{E}_T$  it provides a powerful rejection of the diboson background.

Let us summarize the cuts for the “3-leptons+ $\cancel{E}_T$ ” search:

- $N_l \geq 3$  : at least 3 isolated leptons
- $|M_{l^+l^-} - M_Z| > 6 \text{ GeV}$  : for all possible combinations of  $l^+l^-$  lepton pairs. Implicit: a requirement of a same flavor  $l^+l^-$  pair.
- $\cancel{E}_T$  -cut, two cut values were studied:

$$- \cancel{E}_T > 150 \text{ GeV}$$

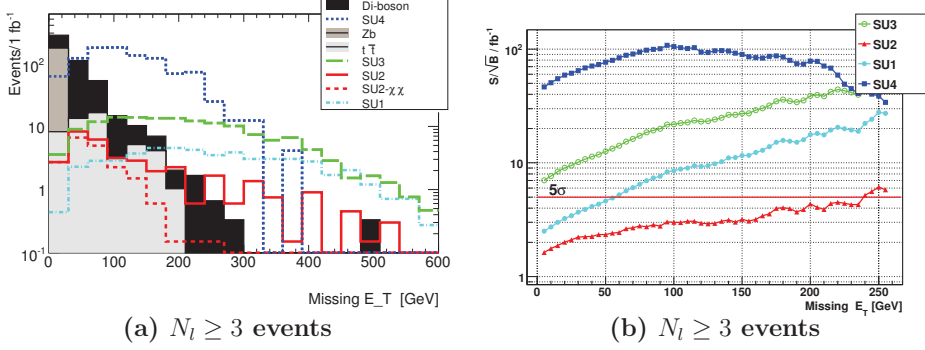


Figure 8.5:  $\cancel{E}_T$  (a) after requiring 3 isolated leptons. Significance  $S/\sqrt{B}$  (b) for different  $\cancel{E}_T$  cuts.

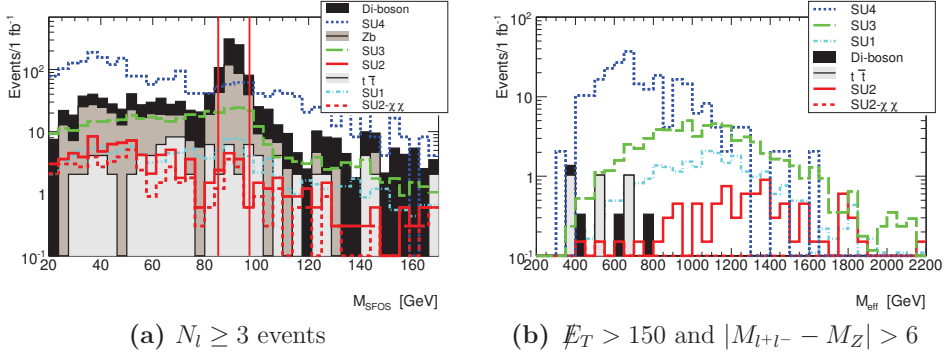


Figure 8.6:  $M_{l+l-}$  (a) after requiring  $N_l \geq 3$  and  $M_{eff}$  (b) after all event selection cut “3-leptons+ $\cancel{E}_T$ ” search.

$$- \cancel{E}_T > 200 \text{ GeV}$$

The analysis is first done without the  $Z$ -mass window cut and the results are presented in Tables 8.4-8.5. The results for the analysis with the  $Z$ -mass window cut can be found in Tables 8.6-8.7.

Comparing the results for the tight cuts,  $p_T^{jet\ 1} > 200$  GeV for “3-leptons+jet” analysis in Table 8.2 and  $\cancel{E}_T > 200$  for “3-leptons+ $\cancel{E}_T$ ” in Table 8.4, we see that the number of background events is similar, while the number of signal events is lower in the “3-leptons+ $\cancel{E}_T$ ” analysis.

Loosening the  $\cancel{E}_T$  cut from 200 to 150 GeV the significances decrease for all benchmark points except SU4 for which the looser cuts are favourable.

Applying the tight cut  $\cancel{E}_T > 200$  GeV and removing events with an invariant mass  $M_{l+l-}$  in the  $Z$ -window, results in almost background free samples, although the number of events is small both for signal and background. This can be seen in Table 8.6. When the numbers become lower than 10 the simple formula  $S/\sqrt{B}$  is no longer a good approximation and the results must be treated with caution.

Figure 8.6 (b) shows the effective mass for trilepton events satisfying the two requirements:  $\cancel{E}_T > 150$  GeV and  $|M_{l+l-} - M_Z| > 6$ . This corresponds to the results of Table

8.7 and one can see that the distributions are dominated by signal.

Results “3-leptons+ $\cancel{E}_T$ ” inclusive analysis using the  $f_{40}(p_T)$ -cut given by Eq. 7.2.  $Z$ -mass window cut is not applied.

| Sample                      | SU1         | SU2 | SU3  | SU4  | Sample                      | SU1                           | SU2 | SU3  | SU4  |
|-----------------------------|-------------|-----|------|------|-----------------------------|-------------------------------|-----|------|------|
| S                           | 33          | 8   | 73   | 147  | S                           | 41                            | 11  | 96   | 314  |
| B                           | 4 (Diboson) |     |      |      | B                           | 12 (5 $t\bar{t}$ , 7 Diboson) |     |      |      |
| $S/B$                       | 9.5         | 2.3 | 20.8 | 42.0 | $S/B$                       | 3.3                           | 0.9 | 7.8  | 25.7 |
| $S/\sqrt{B}$                | 17.7        | 4.4 | 38.9 | 78.5 | $S/\sqrt{B}$                | 11.6                          | 3.2 | 27.3 | 89.8 |
| $S/\sqrt{B + (\sigma_B)^2}$ | 16.6        | 4.1 | 36.4 | 73.5 | $S/\sqrt{B + (\sigma_B)^2}$ | 9.5                           | 2.6 | 22.4 | 73.6 |
| $Z_n$                       | 9           | 3   | 15   | 24   | $Z_n$                       | 7                             | 2   | 13   | 32   |

Table 8.4: Tight  $\cancel{E}_T$  cut:  $\cancel{E}_T > 200$  GeV. Table 8.5: Loose  $\cancel{E}_T$  cut:  $\cancel{E}_T > 150$  GeV.

Results “3-leptons+ $\cancel{E}_T$ ” inclusive analysis using the  $f_{40}(p_T)$ -cut given by Eq. 7.2.  $Z$ -mass window cut  $|M_{l+l^-} - M_Z| > 6$ .

| Sample                      | SU1           | SU2  | SU3   | SU4   | Sample                      | SU1                          | SU2 | SU3  | SU4   |
|-----------------------------|---------------|------|-------|-------|-----------------------------|------------------------------|-----|------|-------|
| S                           | 22            | 6    | 52    | 122   | S                           | 26                           | 9   | 68   | 242   |
| B                           | 0.3 (Diboson) |      |       |       | B                           | 4 (3 $t\bar{t}$ , 1 Diboson) |     |      |       |
| $S/B$                       | 64.9          | 19.0 | 155.1 | 365.2 | $S/B$                       | 6.0                          | 1.9 | 15.3 | 54.6  |
| $S/\sqrt{B}$                | 37.5          | 11.0 | 89.6  | 211.0 | $S/\sqrt{B}$                | 12.6                         | 4.1 | 32.2 | 114.9 |
| $S/\sqrt{B + (\sigma_B)^2}$ | 37.2          | 10.9 | 89.0  | 209.6 | $S/\sqrt{B + (\sigma_B)^2}$ | 11.6                         | 3.8 | 29.7 | 105.9 |
| $Z_n$                       | 12            | 5    | 20    | 33    | $Z_n$                       | 7                            | 3   | 9    | 33    |

Table 8.6: Tight  $\cancel{E}_T$  cut:  
 $\cancel{E}_T > 200$  GeV.

Table 8.7: Loose  $\cancel{E}_T$  cut:  
 $\cancel{E}_T > 150$  GeV.

### 8.2.3 Jet inclusive analysis - summary, conclusions

The discovery potential of two different jet inclusive SUSY searches in trilepton final states has been studied with MC data equivalent of  $1 \text{ fb}^{-1}$ . The first one, the “3-leptons+jet” analysis is based on a requirement of at least one jet with a  $p_T$  of more than 150 GeV or 200 GeV. For both cuts a significance of more than 5 is feasible with  $1 \text{ fb}^{-1}$ . An alternative analysis exploits another typical characteristic of the mSUGRA signal, namely the high  $\cancel{E}_T$ . The “3-leptons+ $\cancel{E}_T$ ” search is investigated requiring  $\cancel{E}_T > 150$  GeV and  $\cancel{E}_T > 200$  GeV. Although the obtained significances are lower than for the analysis based on a tight jet- $p_T$  cut, it is still possible to reach the  $5\sigma$  level within  $1 \text{ fb}^{-1}$  for all studied benchmark points except SU2. Figure 8.7 shows the significance as function of the integrated luminosity for the two inclusive analyses. The presented results are for the looser cuts of  $p_T^{\text{jet1}} > 150$  GeV and  $\cancel{E}_T > 150$  GeV, however both types of searches show promising significances, even with the relatively small amount of data corresponding to the early data taking. With the exception of point SU2, it indicates  $5\sigma$  discovery within approximately  $1\text{--}100 \text{ pb}^{-1}$  for all other points and combinations of event selection cuts. SU2 may reach  $5\sigma$  with around  $1 \text{ fb}^{-1}$ .

The quoted significance  $Z_N$  is calculated using the number of events in tables and therefore with smaller accuracy. It is, however, still interesting as it offers a more correct treatment of the background uncertainty and is known to be a conservative estimate.

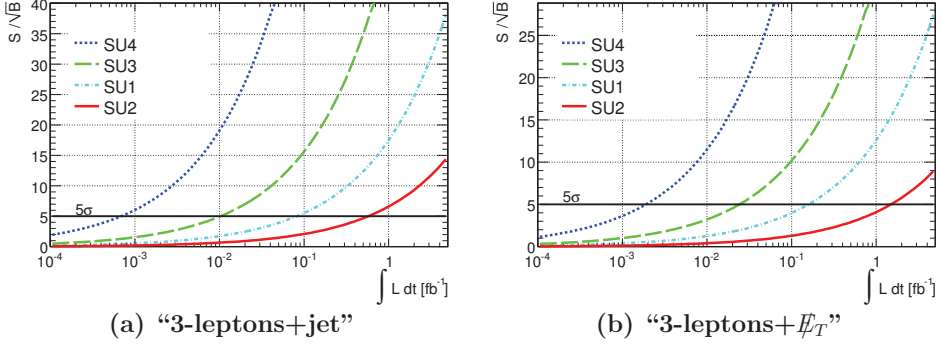


Figure 8.7: Significance  $S/\sqrt{B}$  as function of integrated luminosity for “3-leptons+jet” with requiring  $p_T^{jet1} > 150$  GeV (a) and “3-leptons+ $E_T$ ” requiring  $E_T > 150$  GeV and  $|M_{l+l^-} - M_Z| > 6$  GeV

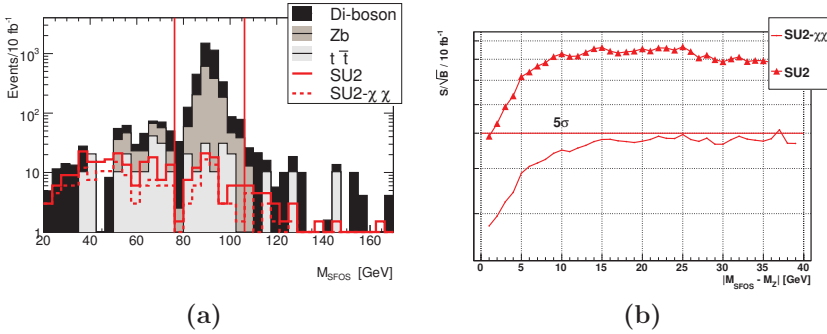


Figure 8.8:  $M_{l+l^-}$  after requiring the presence of a  $l^+l^-$  lepton pair and totally 3 isolated leptons (a). Significance  $S/\sqrt{B}$  as function of the mass window cut  $|M_{l+l^-} - M_Z|$  (b).

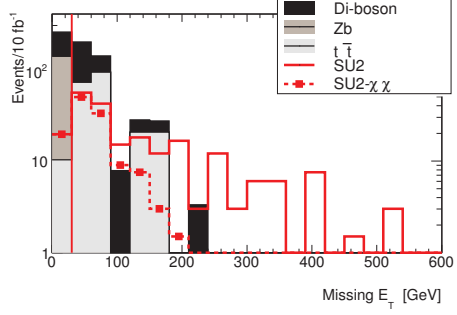
### 8.3 Jet exclusive analysis

The jet exclusive analysis is optimized for direct gaugino production of the Feynman diagram type shown in Figure 6.11 (b), labelled  $SU2 - \chi\chi$ . In SU2, the production of lightest gauginos,  $\chi_2^0\chi_1^\pm$ , has the highest cross section. Gluinos are the lightest strongly interacting SUSY particles in the Focus Point region and pair production of these can have a substantial cross section like in the case of the studied benchmark point SU2. However, these too could be out of the LHC-reach (meaning  $m_{\tilde{g}}$  in the TeV range) leaving the gaugino production as the only accessible SUSY process. The cross section for these processes are relatively low and all distributions in this section are normalized to  $10 \text{ fb}^{-1}$ .

The presence of a same flavour opposite sign lepton pair,  $l^+l^-$ , is fundamental in this analysis and is required from the very beginning along with the presence of three or more well isolated leptons which are selected according to the object definition summarized in Section 8.1.

Other similar trilepton analyses [166], [167] have implemented an additional isolation requirement on the maximum track  $p_T$  in a cone, defined in Section 7.2.2, evaluated for all leptons in an event. The cut is applied on the least isolated lepton in an event requiring

Figure 8.9:  $\cancel{E}_T$  after requiring 3 or more isolated leptons and a  $l^+l^-$  lepton pair with  $|M_{l^+l^-} - M_Z| > 15$  GeV.



$p_{T\max\text{ track}}^{\Delta R=0,X} < 1$  GeV for muons and  $p_{T\max\text{ track}}^{\Delta R=0,X} < 2$  GeV for electrons. In the present analysis which applies a different set of isolation criteria, such a cut did not improve the results in terms of significance and was omitted in order to keep more events.

After the  $l^+l^-$ -pair and trilepton requirement the dominant backgrounds left are  $Zb$  and diboson ( $ZZ, WZ, WW$ ) events. The presence of a  $Z$ -boson suggests to reject events with a  $l^+l^-$  lepton pair whose mass falls within a window around the nominal  $Z$  mass. In case of more than one  $l^+l^-$  lepton combination, the cut is applied on the mass closest to  $M_Z$  in order to enhance the effect of the cut. The distribution of the invariant mass  $M_{l^+l^-}$  of the lepton combinations giving a mass closest to  $M_Z$  is shown in Figure 8.8 (a). Figure 8.8 (b) shows the significance  $S/\sqrt{B}$  as function of the width of the mass window  $|M_{l^+l^-} - M_Z|$ . Based on this plot, the requirement is set to

$$|M_{l^+l^-} - M_Z| > 15 \text{ GeV}. \quad (8.5)$$

The most commonly studied SUSY events starting with squarks or gluinos with subsequent long decay chains have a significantly higher  $\cancel{E}_T$  than SM processes. In direct gaugino production events the situation is different as the escaping LSPs are often back-to-back resulting in a lower  $\cancel{E}_T$ . However, it is still possible to reject SM events by requiring a small minimum amount of  $\cancel{E}_T$ . An optimization study suggests a minimum cut  $\cancel{E}_T > 30$  GeV. The  $\cancel{E}_T$  distribution after  $|M_{l^+l^-} - M_Z| > 15$  GeV requirements is shown in Figure 8.9. The first bin is clearly dominated by the diboson and  $Zb$  events, and the latter are actually totally rejected by the  $\cancel{E}_T$  cut <sup>4</sup>.

In addition to the direct gaugino production ( $\text{SU2-}\chi\chi$ ), point SU2 has also a considerable cross section for gluino pair production. These processes are not considered as SUSY background as the analysis aims at the more general case in which squarks and gluinos may be beyond the LHC energy range and the direct gaugino production is the only SUSY signal. In such processes one does not expect jets with high  $p_T$ , thus it is natural to introduce a jet veto (JV). Figure 8.10 (a) shows the distribution of the leading jet  $p_T$  after the cuts  $|M_{l^+l^-} - M_Z| > 15$  GeV and  $\cancel{E}_T > 30$  GeV are applied, while (b) shows the significance as function of the maximum jet  $p_T$  cut. First when requiring  $p_T^{\text{jet1}}$  less than approximately 120 GeV we can see that the direct gaugino signal ( $\text{SU2-}\chi\chi$ ) is distinguishable from the general SU2 signal which also includes other processes. A jet veto with a cut  $p_T^{\text{jet1}} < 80$  GeV has been chosen and it basically removes the SU2 non-gaugino signal while  $\text{SU2-}\chi\chi$  is reduced by 20%. The  $t\bar{t}$  background is reduced by  $\sim 40\%$  (Table 8.8).

<sup>4</sup>The Monte Carlo sample for the high cross section  $Zb$  process is unfortunately small and the encouraging result needs to be interpreted with some caution.

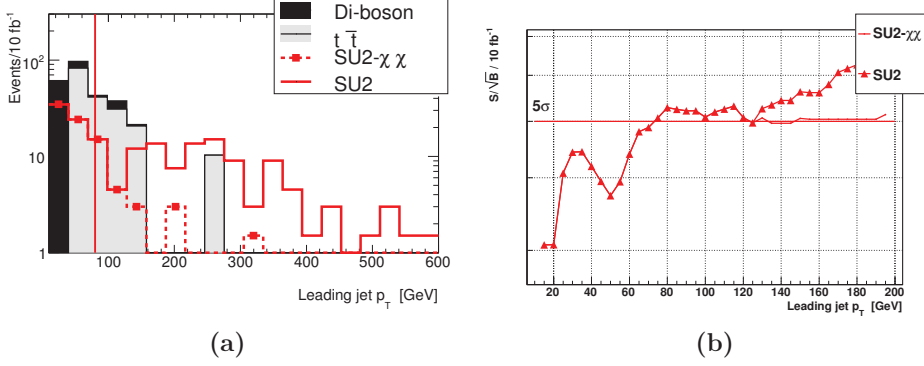
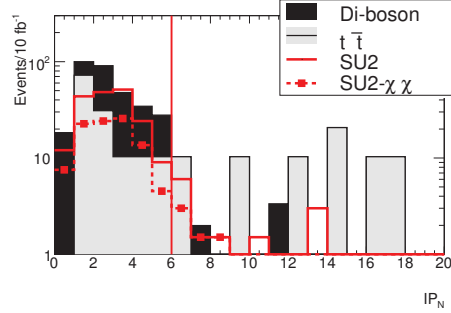


Figure 8.10:  $p_T$  of the leading jet after the  $Z$ -mass window and  $\cancel{E}_T$  cut (a). Significance  $S/\sqrt{B}$  as function of the cut maximum cut on leading jet  $p_T$  (b).

Figure 8.11:  $IP_N$  distribution after the  $\cancel{E}_T$  cut (no jet veto).



After the  $\cancel{E}_T$  cut the largest background consists of  $t\bar{t}$  events which pass the trilepton requirement due to one or more *extra* lepton. The main source of such *extra* leptons is the semi-leptonic  $b$ -decay which occurs at a secondary vertex causing a larger transverse impact parameter  $d_0$  than in case of prompt leptons. Figure 8.11 shows the maximum impact parameter significance  $IP_N = |d_0/\sigma(d_0)|$  per event, after applying all cuts described so far, except the jet veto. A tail consisting mainly of  $t\bar{t}$  events is visible. A cut suggested by the maximum of  $S/\sqrt{B}$  at  $IP_N < 6$  reduces the  $t\bar{t}$  background by  $\sim 35\%$  while only 6 % of the SU2 and SU2- $\chi\chi$  signal is lost. The rejection of the cut is similar whether used after the jet veto or as an alternative to it.

The discovery potential of the analysis has been investigated applying the jet veto *or* the impact parameter requirement as the last cut, as well as a combination of both.

The final list of cuts is as follows:

- $N_l \geq 2$  and at least one same flavour opposite sign pair,  $N_{l+l-} \geq 1$
- $N_l \geq 3$  : 3 isolated leptons
- $Z$ -window :  $|M_{l+l-} - M_Z| > 15$  GeV events with a  $M_{l+l-}$  within the  $Z$ -mass window are rejected
- $\cancel{E}_T > 30$  GeV
- $p_T^{jet1} < 80$  GeV - jet veto



|   | Cut                                 | SU2   | SU2- $\chi\chi$ | $t\bar{t}$ | $Zb$    | WZ     | ZZ    | WW     | $Z\gamma$ | B      |
|---|-------------------------------------|-------|-----------------|------------|---------|--------|-------|--------|-----------|--------|
| 1 | No Cuts                             | 71501 | 64209           | 4448612    | 1593946 | 156675 | 38593 | 401744 | 32867     |        |
| 2 | $N_{l+l-} \geq 1$                   | 1500  | 1116            | 101470     | 578275  | 14875  | 10524 | 11068  | 7330      | 723542 |
| 3 | $N_l \geq 3$                        | 297   | 170             | 320        | 1542    | 1926   | 556   | 16     | 40        | 4400   |
| 4 | $ M_{l+l-} - M_Z  > 15 \text{ GeV}$ | 222   | 124             | 206        | 130     | 260    | 39    | 16     | 10        | 661    |
| 5 | $\cancel{E}_T > 30 \text{ GeV}$     | 202   | 104             | 196        | 0       | 183    | 10    | 16     | 0         | 405    |
| 6 | $p_T^{jet} < 80 \text{ GeV}$ JV     | 86    | 86              | 114        | 0       | 177    | 7     | 16     | 0         | 314    |
| 7 | $IP_N < 6$ with JV                  | 81    | 81              | 72         | 0       | 170    | 6     | 0      | 0         | 248    |
| 8 | $IP_N < 6$ no JV                    | 188   | 98              | 134        | 0       | 177    | 8     | 0      | 0         | 319    |

Table 8.8: Cut-flow table. Number of events corresponds to the number after removal of events with electrons in crack. All numbers are normalized to  $10fb^{-1}$ .

Results for jet exclusive analysis,  $10fb^{-1}$ . Background consists of SM processes.

|   | SU2                                 | $S/B$ | $S/\sqrt{B}$ | $S/\sqrt{B + (\sigma_B)^2}$ | $Z_N$ |
|---|-------------------------------------|-------|--------------|-----------------------------|-------|
| 2 | $N_{l+l-} \geq 1$                   | 0.00  | 1.8          | 0.0                         | N/A   |
| 3 | $N_l \geq 3$                        | 0.1   | 4.5          | 0.3                         | 0     |
| 4 | $ M_{l+l-} - M_Z  > 15 \text{ GeV}$ | 0.3   | 8.6          | 1.7                         | 2     |
| 5 | $\cancel{E}_T > 30 \text{ GeV}$     | 0.5   | 10.0         | 2.4                         | 2     |
| 6 | $p_T^{jet} < 80 \text{ GeV}$ JV     | 0.3   | 4.9          | 1.3                         | 1     |
| 7 | $IP_N < 6$ with JV                  | 0.3   | 5.2          | 1.6                         | 2     |
| 8 | $IP_N < 6$ no JV                    | 0.6   | 10.6         | 2.5                         | 3     |

Table 8.9: All SUSY processes in SU2 are counted as signal.

|   | SU2 $\chi\chi$                      | $S/B$ | $S/\sqrt{B}$ | $S/\sqrt{B + (\sigma_B)^2}$ | $Z_N$ |
|---|-------------------------------------|-------|--------------|-----------------------------|-------|
| 2 | $N_{l+l-} \geq 1$                   | 0.00  | 1.3          | 0.0                         | N/A   |
| 3 | $N_l \geq 3$                        | 0.4   | 2.6          | 0.2                         | 0     |
| 4 | $ M_{l+l-} - M_Z  > 15 \text{ GeV}$ | 0.2   | 4.8          | 0.9                         | 2     |
| 5 | $\cancel{E}_T > 30 \text{ GeV}$     | 0.3   | 5.2          | 1.2                         | 2     |
| 6 | $p_T^{jet} < 80 \text{ GeV}$ JV     | 0.3   | 4.9          | 1.3                         | 1     |
| 7 | $IP_N < 6$ with JV                  | 0.3   | 5.2          | 1.6                         | 2     |
| 8 | $IP_N < 6$ no JV                    | 0.3   | 5.6          | 1.5                         | 3     |

Table 8.10: Signal is defined as direct gaugino production only.

- $IP_N < 6$ ,  $IP_N$  is the maximum per event

The cut flow, Table 8.8, for signal and background shows in row 6 the number of events after the jet veto and in row 7 the results after both the jet veto and the cut on the impact parameter. Row 8 shows the number of events after the impact parameter cut, without the jet veto.

### 8.3.1 Jet exclusive analysis - summary, conclusions

The resulting significances for the exclusive SU2 signal which also includes other processes than direct gaugino production are presented in Table 8.9, while Table 8.10 shows the results for SU2- $\chi\chi$ , the direct gaugino production.

The significance  $S/\sqrt{B}$  with  $10fb^{-1}$  is for the direct gaugino production 4.85 after the jet veto and 5.16 with the impact parameter cut (Table 8.9, row 6 and 7). Without the jet veto but with the impact parameter cut, the significance is 5.59 (Table 8.10, row 8).

The significance for the inclusive SU2 signal after the jet veto and the impact parameter cut are the same as those quoted above for the direct gaugino production because these

are only SUSY processes surviving the jet veto. Without the jet veto, but still with the impact parameter cut a significance of 10.55 is obtained for  $10 \text{ fb}^{-1}$  (Table 8.9, row 8). The large increase as compared to results with the jet veto are due to gluino pair production events.

The discovery reach in terms of the statistical significance  $S/\sqrt{B}$  is shown in Figure 8.12 as a function of the integrated luminosity.

For the inclusive SU2 signal, without the jet veto, a  $5\sigma$  discovery signal is reached after about  $3 \text{ fb}^{-1}$ . While the direct gaugino production, with and without the jet veto, comes close to  $5\sigma$  discovery after approximately  $10 \text{ fb}^{-1}$ .

The more conservative significance measures which include background uncertainty,  $S/\sqrt{B + (\sigma_B)^2}$  and  $Z_N$  give both values below the discovery limit with  $10 \text{ fb}^{-1}$ .

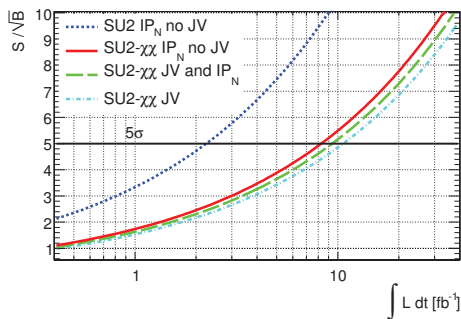


Figure 8.12: The discovery reach in terms of the statistical significance  $S/\sqrt{B}$  as function of the integrated luminosity.

Clearly, such process may be difficult to study during early data taking, not only because of the low cross sections, but they will also require good understanding of the detector, as well as different effects like pile-up when running at high luminosities ( $> 10 \text{ fb}^{-1}/\text{year}$ ).

### 8.3.2 Trigger efficiency for 3-lepton events

Both the jet inclusive and exclusive analyses depend on an efficient lepton trigger. Although the applied lepton  $p_T$  threshold of 10 GeV is considered low, we observe in Figure 8.2 that the  $p_T$  of the two leading leptons in SUSY trilepton events is rather high providing relatively high trigger efficiency. A brief study of the trigger or combination of isolated lepton triggers at Event Filter level, `PassedEF_e25i`  $\cup$  `Passed_mu20i`, shows for events with at least three isolated leptons an efficiency of 97% for all studied SUSY samples, except SU2 which has an efficiency of 95%. The  $Zb$  and  $t\bar{t}$  backgrounds have a slightly lower efficiency of 93% while it is around 97% or better for the diboson backgrounds. The significances for jet inclusive analysis calculated after requiring the events to pass the above described trigger requirements are roughly 3% lower for all SUx points.

## 8.4 Preliminary Conclusions and Outlook

The interesting class of SUSY final states with three leptons requires both an efficient selection of well isolated leptons, as well as strong rejection of *extra* leptons often coming

from a jet. The first part of this analysis focused therefore on different isolation variables and additional requirements on leptons defined to be inside a jet. The isolation cuts were optimized for trilepton events and the single lepton selection efficiency and fake rates were studied. The best results were obtained applying the relative track-based isolation with an additional cut for leptons in jets consisting of a linear cut which combines a higher  $p_T$  threshold with a tight isolation.

After applying these preselection requirements two modes of trilepton analysis were studied: a jet inclusive and a jet exclusive. The jet inclusive channel shows a promising discovery potential for most SUSY benchmark points even with an integrated luminosity of only  $1\text{fb}^{-1}$  which is considered achievable during early data taking. The jet exclusive analysis focuses on the direct gaugino production. Due to both a low cross section and more complex event selection, it will be feasible only with higher integrated luminosities and good detector understanding. A  $5\sigma$  signal may be reached with approximately  $10\text{fb}^{-1}$ .

Compared to similar analyses like those presented in the CSC book [78], the significances obtained for both types of analysis improve with the suggested set of isolation criteria as the isolation requirements are directly related to suppression of some of the dominant backgrounds.

We have also seen that for some sets of event selection cuts one will have to deal with small numbers of both signal and background. Such situation will require more sophisticated statistical treatment than the simple measures used in this analysis. In addition, the main significance measure of this analysis  $S/\sqrt{B}$  does not take into account background uncertainties and is known to give too optimistic results. The choice of significance measure has also an impact on the cut optimization.

Some of the official ATLAS 14 TeV MC simulations of SM backgrounds could not be used due to cuts applied at the event generation level which were incompatible with the trilepton event selection. Although the used backgrounds represent the dominant processes, the list will have to be expanded in order to obtain a more realistic picture. Chapter 9 will present SUSY searches based on a new MC simulation at 10 TeV which corresponds to the first revised LHC plan defined after the 19th September (2008) magnet incident. Many crucial SM samples were simulated without strict event filters and were no longer inconsistent with the trilepton event selection. Despite the lower center of mass energy, the 10 TeV study will show a more complete picture of the SM backgrounds.

The MC simulation used in the 14 TeV analysis is based on ATLAS release 12. Since the CSC MC production, ATLAS software has been through considerable development, and the performance of the particle reconstruction has been improved. The ATLAS jet expert group recommends also to move to more reliable jet reconstruction algorithms than the seeded cone algorithm.

An obvious weakness of the analyses presented in this thesis is that they are based on MC truth information. As the data starts arriving, the lepton reconstruction performance and trigger efficiencies will have to be derived using data driven methods. All quantities used in the analysis will have to be well understood. In addition, the analysis will require careful study of different effects like pile-up and beam-gas collisions which are expected to influence the reconstruction performance.

ATLAS recorded its first collision data at  $\sqrt{s} = 900\text{ GeV}$  during the last two months of 2009. Although the energy and integrated luminosity is too low to perform any detailed performance study, it is still very interesting to get an idea of how well MC reproduces the data. Chapter 11 will therefore present a brief study of these first collision data focusing

on physics objects used in the MC based analyses. Moreover, such simple comparison studies can give a very preliminary indication of the understanding of the detector and reliability of MC studies. The study includes also a suggestion for a data-driven study of electron reconstruction.

# Chapter 9

## Trilepton SUSY searches with early data at 10 TeV

After the coupling failure on the 19th of September 2008, the LHC accelerator team in collaboration with the experiments decided on a new start-up plan. During the first months of 2009 it became clear that the center of mass energy had to be reduced to 10 TeV in order to find a compromise between the machine security and the physics and discovery potentials [76]. The new plans triggered a large centralized MC production at 10 TeV which became the basis for the preparations of the early data-taking studies. This chapter presents trilepton search strategies using these MC samples at 10 TeV. Meanwhile, the plans have been updated. In the beginning of 2010 it was decided to run LHC for approximately two years at the center of mass energy of 7 TeV [77].

This chapter will address many of the shortcomings of the earlier presented 14 TeV analysis. First of all it includes an important range of new SM backgrounds which are simulated with latest versions of the ATLAS software. The infrared safe Anti- $K_T$  jets will be introduced along with a proposal for different kinematical cuts on muons within a jet cone. The preparations for early data analysis at lower center of mass energies focus on the low mass point SU4 and jet inclusive analysis. The new phenomenological picture called for a revision of event selection and a cut optimization, however, this time based on the significance measure  $Z_N$  which includes background uncertainty. This is particularly important during early analysis as it is expected to be a large systematic effect.

In Chapter 10 the strategies developed for the mSUGRA benchmark points will be applied on a broad set of points within the less restricted MSSM24 model.

### 9.1 mSUGRA benchmark points at 10 TeV

For the new round of simulation at 10 TeV ATLAS decided not to generate the SU2 (Focus Point) sample. This simulation focused on early searches feasible with a sample of  $200 \text{ pb}^{-1}$ , a priority which clearly excluded SU2 due to its low cross section.

The mSUGRA parameters defining the points which were simulated were unchanged as compared to the 14 TeV samples (Table 6.1). The difference is the lower center of mass energy which leads to reduced cross sections. For the two benchmark points with highest cross sections, SU4 and SU3, the reduction factor is 4 and 5, respectively. In addition, due to the energy dependent parton distributions, the rates of the different production processes are slightly changed as shown in Table 9.1 for SU3 and in Table 9.2 for SU4.

| 14 TeV   |                     |      | 10 TeV   |                     |      |
|--|---------------------|------|--|---------------------|------|
|  | $\sigma_{NLO}$ [fb] | [%]  |  | $\sigma_{NLO}$ [fb] | [%]  |
| Total  | 27700.0             |      | Total  | 5460.0              |      |
| $qg \rightarrow \tilde{q}_{R/L}\tilde{g}$                        | 12287.8             | 44.4 | $qg \rightarrow \tilde{q}_{R/L}\tilde{g}$                        | 2085.9              | 38.2 |
| $qq \rightarrow \tilde{q}_L\tilde{q}_L$                          | 3788.0              | 13.7 | $qq \rightarrow \tilde{q}_L\tilde{q}_L$                          | 806.5               | 14.8 |
| $qq \rightarrow \tilde{q}_L\tilde{q}_R + \tilde{q}_R\tilde{q}_L$ | 3282.4              | 11.8 | $qq \rightarrow \tilde{q}_R\tilde{q}_R$                          | 697.3               | 12.8 |
| $qq \rightarrow \tilde{q}_R\tilde{q}_R$                          | 2993.0              | 10.8 | $qq \rightarrow \tilde{q}_L\tilde{q}_R + \tilde{q}_R\tilde{q}_L$ | 672.2               | 12.3 |
| $gg \rightarrow \tilde{g}\tilde{g}$                              | 2108.0              | 7.6  | $qq \rightarrow \tilde{\chi}_1^+\tilde{\chi}_2^0$                | 279.0               | 5.1  |
| $qq \rightarrow \tilde{\chi}_1^+\tilde{\chi}_2^0$                | 754.8               | 2.7  | $gg \rightarrow \tilde{g}\tilde{g}$                              | 252.3               | 4.6  |
| $qg \rightarrow \tilde{q}_L\tilde{\chi}_1^+$                     | 671.7               | 2.4  | $qq \rightarrow \tilde{\chi}_1^+\tilde{\chi}_1^+$                | 150.7               | 2.8  |

Table 9.1: Comparison of cross sections for the dominant production processes for the SU3 benchmark point at 14 TeV and at 10 TeV.

| 14 TeV   |                     |      | 10 TeV   |                     |      |
|--|---------------------|------|--|---------------------|------|
|  | $\sigma_{NLO}$ [fb] | [%]  |  | $\sigma_{NLO}$ [fb] | [%]  |
| Total  | 403480.0            |      | Total  | 107400.0            |      |
| $qg \rightarrow \tilde{q}_{R/L}\tilde{g}$                        | 180960.7            | 44.9 | $qg \rightarrow \tilde{q}_{R/L}\tilde{g}$                        | 43717.7             | 40.7 |
| $qq \rightarrow \tilde{q}_L\tilde{q}_L$                          | 76217.4             | 18.9 | $qq \rightarrow \tilde{q}_L\tilde{q}_L$                          | 25292.7             | 23.6 |
| $gg \rightarrow \tilde{g}\tilde{g}$                              | 67118.9             | 16.6 | $gg \rightarrow \tilde{g}\tilde{g}$                              | 13086.7             | 12.2 |
| $qq \rightarrow \tilde{q}_L\tilde{q}_R + \tilde{q}_R\tilde{q}_L$ | 25076.3             | 6.2  | $qq \rightarrow \tilde{q}_L\tilde{q}_R + \tilde{q}_R\tilde{q}_L$ | 7276.4              | 6.8  |
| $qq \rightarrow \tilde{q}_R\tilde{q}_R$                          | 24672.8             | 6.1  | $qq \rightarrow \tilde{q}_R\tilde{q}_R$                          | 6632.0              | 6.2  |
| $qq \rightarrow \tilde{\chi}_1^+\tilde{\chi}_2^0$                | 9744.0              | 2.4  | $qq \rightarrow \tilde{\chi}_1^+\tilde{\chi}_2^0$                | 4473.2              | 4.2  |
| $qg \rightarrow \tilde{q}_L\tilde{\chi}_1^+$                     | 5789.9              | 1.4  | $qq \rightarrow \tilde{\chi}_1^+\tilde{\chi}_1^+$                | 2389.7              | 2.2  |

Table 9.2: Comparison of cross sections for the dominant production processes for the SU4 benchmark point at 14 TeV and at 10 TeV.

The general tendency is that the fractions of the total cross section for processes involving gluons in the initial state and gluinos among the produced particles are reduced, while quark interactions that produce squarks and gauginos have increased cross section. In both SU3 and SU4, the gaugino pair production with the highest cross section is for the wino like pair of  $\tilde{\chi}_1^\pm\tilde{\chi}_2^0$  which in turn can lead to a jet exclusive trilepton signal. At 10 TeV the fraction of the total cross section for these processes is nearly doubled. However, the cross sections for direct gaugino production are very small and priority is given to the jet inclusive searches.

According to the new plans from January 2010, LHC will start at the even lower energy of 7 TeV which implies further reduction of the cross sections (LO:  $\sigma_{SU4} = 48.0$  pb,  $\sigma_{SU3} = 1.8$  pb). Focusing on searches for deviations from the SM expectations, the most relevant point will still be SU4.

## 9.2 Standard Model background samples at 10 TeV

One of the weaknesses of the 14 TeV trilepton analysis presented in Chapter 8 was the limited range of SM backgrounds. Some of the relevant processes were simulated centrally, however the applied filters were inconsistent with the trilepton event selection. Although

| Process                 | Event Generator | $\sigma_{LO}$ [pb] | $\sigma_{NLO}$ [pb] | Total number of events |
|-------------------------|-----------------|--------------------|---------------------|------------------------|
| $Z \rightarrow ll + Np$ | ALPGEN          | 3621.3             | 4418.0              | $1.9 \times 10^6$      |
| $W \rightarrow l + Np$  | ALPGEN          | 39719.3            | 48457.5             | $7.7 \times 10^6$      |
| $Wbb$                   | ALPGEN          | 14.6               | 17.8                | $4.5 \times 10^4$      |
| $Zbb$                   | ALPGEN          | 60.3               | 73.6                | $1.5 \times 10^6$      |
| $bb$                    | ALPGEN          | 464025.0           | 566110.5            | $5.5 \times 10^5$      |
| $WW$                    | MC@NLO          | 7.1                | 7.1                 | $2.0 \times 10^5$      |
| $ZZ$                    | MC@NLO          | 18.4               | 18.4                | $4.0 \times 10^4$      |
| $WZ$                    | MC@NLO          | 59.0               | 59.0                | $4.5 \times 10^4$      |
| $W\gamma$               | Pythia          | 22.4               | 29.1                | $2.5 \times 10^4$      |
| $Z\gamma$               | Pythia          | 11.1               | 14.4                | $2.5 \times 10^4$      |
| $t\bar{t}$              | MC@NLO          | 205.5              | 205.5               | $2.0 \times 10^6$      |
| $Wt$                    | AcerMC          | 14.4               | 20.2                | $1.0 \times 10^3$      |
| Single top $t$ -chan.   | AcerMC          | 41.1               | 40.3                | $3.0 \times 10^4$      |
| SU1                     | HERWIG          | 2.4                | 3.2                 | $1.0 \times 10^4$      |
| SU3                     | HERWIG          | 5.5                | 7.3                 | $1.5 \times 10^4$      |
| SU4                     | HERWIG          | 107.4              | 147.1               | $1.7 \times 10^5$      |

Table 9.3: Summary of SM background and SUSY signal samples 10 TeV. Cross section are from AMI database [126], K-factors for calculation of  $\sigma_{NLO}$  from [78, 170]

the  $Zb$  process accounts for one of the largest backgrounds, it could only give a rough estimate of the  $Z + jets$  background. These filters were not applied in the new MC production at 10 TeV, allowing for example to include in the study the more detailed ALPGEN [168] samples of leptonically decaying gauge bosons accompanied by 0 to 5 partons. Leptonically decaying gauge bosons produced in association with two  $b$ -quarks and 0 to 3 partons were also simulated using ALPGEN. The diboson samples  $WW$ ,  $ZZ$  and  $WZ$  were generated with MC@NLO [148, 149] and split up in different decay modes. While the light quark di-jet background is strongly reduced by the tripleton requirements, the production of  $b$ -quark pairs with its very high cross section and leptons from decay of B-mesons that may pass the isolation cuts is a potentially difficult background [169]. This study includes therefore ALPGEN samples for  $bb$ -production in association with 0 to 4 partons. However, the simulated integrated luminosity is very low. Table 9.3 gives a summary of the various SM backgrounds and SUSY signal samples used in the present study. A full detailed list is given in Appendix C.

Next to leading order (NLO) cross sections are obtained via K-factors. Originally these were calculated for the 14 TeV analysis, however, they are not expected to change drastically for the planed lower center of mass energies [170].

### 9.3 Object definition

The object definition applied in this study follows the general recommendation to build on the experience from the CSC analysis and apply simple standard definitions. While these may not be optimal, one should keep in mind that it is only an intermediate study. All features of the physics objects will in the close future be carefully studied with real

data. The basic selection criteria will therefore be the same as for most SUSY analyses at 14 TeV. A simple alternative to the lepton jet separation criterion will be presented.

The definitions of the relevant physics objects, electrons, muons, jets and  $\cancel{E}_T$  are very similar as for the 14 TeV analysis in Chapter 8, Section 8.1. The differences are:

### Electrons and Muons

- Isolation:  $E_{\text{cal}}^{\Delta R=0.2} < 10 \text{ GeV}$
- Lepton-jet separation. Three different options will be discussed.

### Jets

- The default choice has so far been the cone algorithm with  $\Delta R = 0.4$  (`Cone4H1Tower` container). This analysis will introduce the Anti- $K_T$  algorithm [171] (`AntiKtH1Tower` container).

## 9.3.1 Lepton jet separation

After selecting the physics objects it is common in many analyses to require that no lepton should be closer to a jet than  $\Delta R = 0.4$ . This will be referred to as the  $\Delta R(l, j) > 0.4$  requirement.

The performance of the reconstruction and the applied object definition can be measured by looking at the single lepton efficiency defined in Eq. 7.6, page 153.

The fake rate is defined as:

$$\mathcal{F}_\ell \equiv \frac{n_\ell^{\text{Reco not-matched to MC or from jet}}}{n_{\text{jet}}^{\text{MC}}}, \quad (9.1)$$

where the numerator includes both reconstructed leptons which are *not-matched* to a generator-level lepton and those that do have a MC match, but which are from jets, (*non-prompt*).

The definition of the fake rate is different from the one in Section 7.5 as it focuses on leptons originating from SUSY,  $Z$ ,  $W$  or  $\tau$ . This definition shows better the effect of requirements that aim at rejecting such *non-prompt* leptons.

As shown in Section 7.2.3, the lepton-jet separation requirement is very efficient in rejecting the *non-prompt* muons. However, it also reduces strongly the muon efficiency. From Figures 7.18 and 7.8 which show the distribution of  $\Delta R(l, j)$  for electrons and muons respectively, it is clear that the effect of a separation cut of 0.4 is small for electrons due to the electron-jet overlap removal. For muons, on the other hand, it is interesting to study alternative requirements which combine the proximity of the *non-prompt* muons to jet with an additional cut on a variable that can discriminate between *prompt* and *extra* muons. In Section 7.2.3 it was shown that a tight  $p_T$  dependent isolation cut (Eq. 7.2) is capable of reducing the efficiency loss while at the same time keeping a low fake rate. Since this analysis aims at early data taking it should not rely too much on fine tuned MC based cuts. However, the basic features of leptons from the various sources are valid. The idea is therefore to apply a simple cut which can be relatively reliable also for early analysis. The  $\Delta R(l, j)$  requirement itself has the advantage that it is purely geometrical and less sensitive to shortcomings of the MC simulation of the detector response [172]. A characteristic feature of the *extra* leptons is that they are relatively soft. Figure 9.1



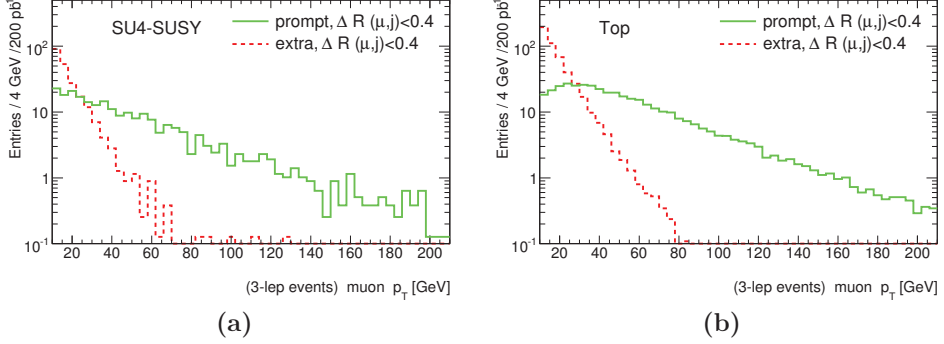


Figure 9.1:  $p_T$  for muons which are closer to a jet than  $\Delta R = 0.4$ . Histograms show muons in trilepton events in SU4 (a) and  $t\bar{t}$  (b) samples. (Cone tower jets  $\Delta R = 0.4$ .)

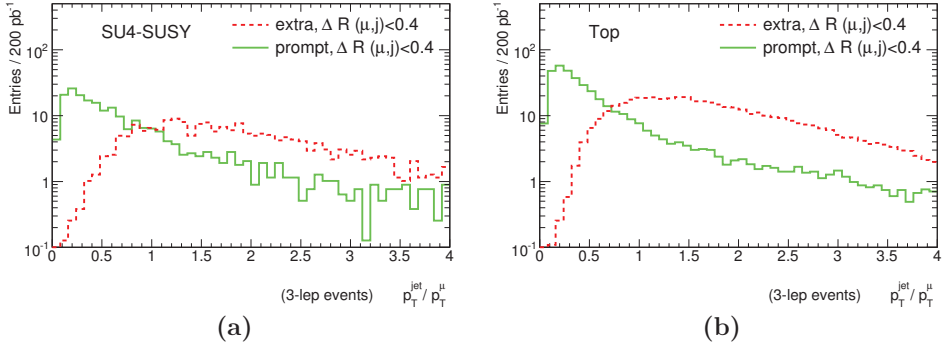


Figure 9.2:  $p_T^{jet}/p_T^\mu$  for muons which are closer to a jet than  $\Delta R = 0.4$ . Histograms show muons in trilepton events in SU4 (a) and  $t\bar{t}$  (b) samples. (Cone tower jets  $\Delta R = 0.4$ .)

shows the  $p_T$  distribution of muons within a jet ( $\Delta R(\mu, j) < 0.4$ ) in trilepton events. Based on the trilepton study at 14 TeV presented in the previous chapter, keeping muons close to a jet only if  $p_T^\mu > 40$  GeV rejects most *extra* muons while at the same time reduces the loss of *prompt* ones. The cut could of course be optimized, however without a precise knowledge of real data, the choice of such a cut is guided by simplicity and strong background rejection.

Another characteristic feature of muons<sup>1</sup> close to a jet, in QCD processes or in SUSY cascade decays, is that in most cases  $p_T^\mu > p_T^{jet}$ . In heavy flavor decays or in punch-through the situation tends to be the opposite,  $p_T^\mu < p_T^{jet}$ , even if the muon often carries a significant fraction of the jet momentum. The ratio of momentum of the jet found closest to the muon, and the muon  $p_T^{jet}/p_T^\mu$  is therefore also a good candidate for a discriminating variable [172]. Figure 9.2 shows the distribution of the  $p_T^{jet}/p_T^\mu$  for *extra* and *prompt* muons within a jet ( $\Delta R(\mu, j) < 0.4$ ) in SU4 and  $t\bar{t}$  events with three or more leptons. Following the suggestion from the ATLAS Top working group study, a cut is set at  $p_T^{jet}/p_T^\mu < 0.5$ . This cut is not optimized for SUSY and is meant more as a proof of concept study.

Table 9.4 summarizes the single lepton efficiencies and fake rates for the three require-

<sup>1</sup>The observation is valid for leptons in general, however here muons are in focus.

| Efficiency % |                        |                      |                                     | Fake rate $\times 10^{-3}$ |                      |                                     |
|--------------|------------------------|----------------------|-------------------------------------|----------------------------|----------------------|-------------------------------------|
| Electrons    |                        |                      |                                     |                            |                      |                                     |
|              | $\Delta R(l, j) > 0.4$ | $\parallel p_T > 40$ |                                     | $\Delta R(l, j) > 0.4$     | $\parallel p_T > 40$ |                                     |
| SU4          | $74.4 \pm 0.6$         | $76.3 \pm 0.6$       | $81.2 \pm 0.6$                      | $2.60 \pm 0.06$            | $2.62 \pm 0.06$      | $2.93 \pm 0.07$                     |
| $t\bar{t}$   | $81.3 \pm 0.5$         | $82.6 \pm 0.5$       | $85.5 \pm 0.5$                      | $3.60 \pm 0.08$            | $3.63 \pm 0.08$      | $3.82 \pm 0.08$                     |
| Muons        |                        |                      |                                     |                            |                      |                                     |
|              | $\Delta R(l, j) > 0.4$ | $\parallel p_T > 40$ | $\parallel p_T^{jet}/p_T^\mu < 0.5$ | $\Delta R(l, j) > 0.4$     | $\parallel p_T > 40$ | $\parallel p_T^{jet}/p_T^\mu < 0.5$ |
| SU4          | $70.7 \pm 0.6$         | $79.3 \pm 0.6$       | $78.9 \pm 0.6$                      | $0.65 \pm 0.03$            | $0.82 \pm 0.04$      | $0.73 \pm 0.03$                     |
| $t\bar{t}$   | $75.6 \pm 0.4$         | $84.6 \pm 0.5$       | $84.9 \pm 0.5$                      | $0.66 \pm 0.03$            | $0.88 \pm 0.04$      | $0.76 \pm 0.04$                     |

Table 9.4: Efficiencies and fake rates for combinations of lepton jet separation  $\Delta R(l, j) > 0.4$  and additional kinematical cuts. The  $\parallel$  sign means logical OR combination of the lepton jet separation requirement and  $p_T > 40$  (9.3) and  $p_T^{jet}/p_T^\mu < 0.5$  (9.4). Fixed cone jets are used.

ments:

$$\Delta R(l, j) > 0.4 \text{ only} \quad (9.2)$$

$$\Delta R(l, j) > 0.4 \parallel p_T > 40 \quad (9.3)$$

$$\Delta R(l, j) > 0.4 \parallel \frac{p_T^{jet}}{p_T^\mu} < 0.5. \quad (9.4)$$

The sign  $\parallel$  stands for a logical OR combination of the two cuts. Due to the electron jet overlap removal there are only few electrons within  $\Delta R = 0.4$ . The effect of the lepton-jet separation on electrons is therefore very small and the cut in Eq. 9.4 it was not applied for them.

For muons on the other hand, the gain in efficiency is about 10% when adding a kinematical cut to the muon-jet separation requirement, while the fake rates stay at the same level. The performance of the simple  $p_T$  cut (9.3) is comparable to the  $p_T^{jet}/p_T^\mu$  cut (9.4), although the last gives lower muon fake rate.

### 9.3.2 Anti- $K_T$ Jets

So far in the analysis the default jet reconstruction algorithm has been the iterative seeded fixed  $\eta - \phi$  cone algorithm with  $\Delta R = 0.4$ . The cone jet reconstruction algorithm was described in Section 7.1

Although the cone algorithm has been the most used in ATLAS so far, it is known not to be an infrared and collinear safe and therefore not considered as the best choice in hadron collisions. Such events are much more complicated than clean  $e^+e^-$  collisions where it is possible to work with exclusive jet cross sections. In a hadron-hadron interaction only a fraction of the hadrons in the final state are associated with the hard jets. The accompanying hadronic activity comes from the underlying event and particles with low transverse momentum, but possibly large momenta collinear with the beam axis [173].

The jet performance group in ATLAS has therefore recommended to move to the Anti- $K_T$  algorithm which is more appropriate in a hadronic collider environment. The  $K_T$  algorithm family groups protojets which are nearby in the momentum space defined by the distance  $d_i = p_{Ti}$ . For each pair of protojets a relative distance is defined as

$$d_{ij} = \min(p_{Ti}, p_{Tj}) \Delta R^2(ij) / R \quad (9.5)$$

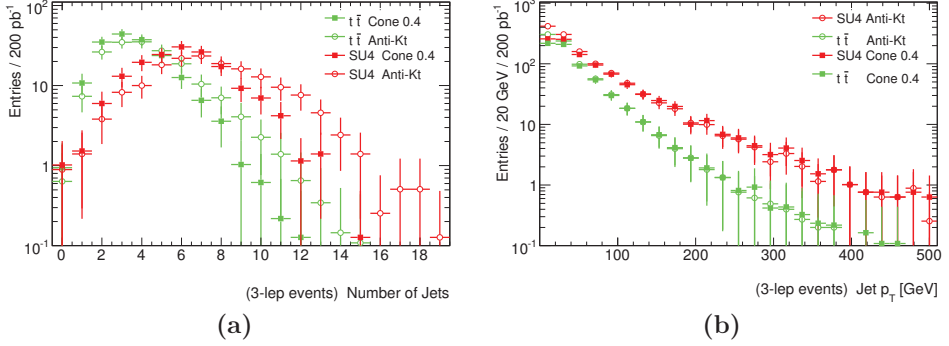
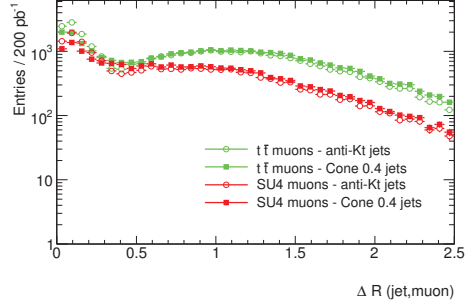


Figure 9.3: Comparison of the number of jets (a) and jet  $p_T$  (b) in SU4 and  $t\bar{t}$  samples for tripleton events as reconstructed by the Anti- $K_T$  and fixed cone  $\Delta R = 0.4$  algorithms. The plots show tripleton events where the leptons close to a jet have not been subject to any cut.

Figure 9.4: Comparison of  $\Delta R$  between muons and the closest jet as reconstructed by the Anti- $K_T$  and fixed cone  $\Delta R = 0.4$  algorithms. All events.



where

$$\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2. \quad (9.6)$$

The parameter  $R$  is the jet radius resolution corresponding to the cone size in a cone algorithm. Next, the  $d_{min} = \min(d_i, d_{ij})$  is searched for. If  $d_{min} = d_i$  the protojet is “not mergeable” and it is defined as a jet and removed from the list of protojets. In case  $d_{min} = d_{ij}$ , the two are merged according to a certain schema (the default is a sum of the 4-momenta of the two protojets). This minimization procedure is repeated until there are no protojets left. The Anti- $K_T$  algorithm follows the same procedure but with a redefinition of the distance  $d_i = 1/p_{Ti}$  so that the most energetic jets are clustered first.

This analysis uses the Anti- $K_T$  tower jets which became available in the official AODs from ATLAS release 15 reconstruction. It is a known feature that the Anti- $K_T$  algorithm gives a high number of jets where the least energetic ones are expected to account for the additional jet activity not related to the hard process [172, 173]. Figure 9.3 shows, for SU4 and  $t\bar{t}$ , a comparison of the number of jets (a) and the jet  $p_T$  (b) as reconstructed by the fixed cone and Anti- $K_T$  algorithms, both with  $\Delta R = 0.4$ , in tripleton events. The histograms confirm the expectation of a higher number of jets and softer  $p_T$  spectrum for Anti- $K_T$  jets. As a consequence of the higher number of jets, one finds a larger fraction of muons close to a jet than with the cone algorithm. This is demonstrated in Figure 9.4 which shows the  $\Delta R$  distance between the muon and the closest jet as given by the two

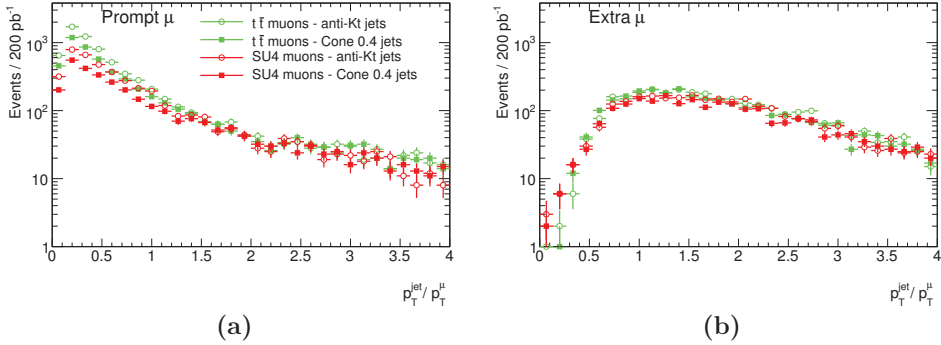


Figure 9.5: Comparison of the relative  $p_T^{jet}/p_T^\mu$  for *prompt* muons (a) and *extra* muons (b) within  $\Delta R(\mu, j) = 0.4$ , for SU4 and  $t\bar{t}$  events as reconstructed by the Anti- $K_T$  and fixed cone  $\Delta R = 0.4$  algorithms. All events.

jet algorithms (for all events, not only trilepton events).

Considering the proposed kinematical cuts, Eq. 9.3 and 9.4, the muon  $p_T$  distribution is not affected by the choice of jet algorithm. The only difference is that Anti- $K_T$  leads to a relative increase of “close-to-a-jet” muons. The relative momentum  $p_T^{jet}/p_T^\mu$ , on the other hand, is sensitive to the softer spectrum of the Anti- $K_T$  jets and gives a better separation between the *extra* and *prompt* muons with higher  $p_T$ . This effect is visible from the  $p_T^{jet}/p_T^\mu$  distributions in Figure 9.5 (a) for *prompt* muons and for the *extra* muons in Figure 9.5 (b), all within  $\Delta R(\mu, j) = 0.4$ , for SU4 and  $t\bar{t}$  (all events).

Table 9.5 gives a summary of the single lepton efficiencies and fake rates for the same combinations of cuts as in Table 9.4, but now using the Anti- $K_T$  jets. Comparing Table 9.4 and 9.5 for electrons we observe a small gain in efficiency for the simple rejection of electrons failing the  $\Delta R(e, j) > 0.4$  separation and when applying it together with the  $p_T > 40$  GeV cut. The fake rates are slightly higher with Anti- $K_T$  jets. The rightmost column of the table where no cuts are applied on electrons close to a jet is unchanged.

In case of muons, the  $\Delta R(\mu, j) > 0.4$  requirement with Anti- $K_T$  jets reduces strongly the muon efficiency as there are more jets to be vetoed against. At the same time the fake rates are significantly lower than for the cone jets. Applying the additional kinematical cuts  $p_T > 40$  GeV or  $p_T^{jet}/p_T^\mu < 0.5$  improves the muon efficiency while still providing lowest fake rates.

In conclusion, using the Anti- $K_T$  jets and requiring muon-jet separation, the efficiency may be greatly improved by combining it with an additional cut like proposed in Equations 9.3 and 9.4. While both  $p_T^\mu > 40$  GeV and  $p_T^{jet}/p_T^\mu < 0.5$  improve the efficiency, there is a hint that the latter gives the lowest fake rate.

As mentioned, this is only a “proof of concept” study, but the observation is interesting. Optimizing the lepton isolation towards high efficiency and low fake rate is important in order to reduce the background due to *extra* leptons, especially from heavy flavor jets like in top events.

| Efficiency % |                        |                        |                                       | Fake rate $\times 10^{-3}$ |                        |                                       |
|--------------|------------------------|------------------------|---------------------------------------|----------------------------|------------------------|---------------------------------------|
| Electrons    |                        |                        |                                       |                            |                        |                                       |
|              | $\Delta R(l, j) > 0.4$ | $\parallel_{p_T > 40}$ |                                       | $\Delta R(l, j) > 0.4$     | $\parallel_{p_T > 40}$ |                                       |
| SU4          | $77.3 \pm 0.6$         | $77.7 \pm 0.6$         | $81.2 \pm 0.6$                        | $2.70 \pm 0.06$            | $2.71 \pm 0.06$        | $2.83 \pm 0.07$                       |
| $t\bar{t}$   | $83.4 \pm 0.4$         | $83.8 \pm 0.5$         | $85.4 \pm 0.5$                        | $3.70 \pm 0.08$            | $3.71 \pm 0.08$        | $3.82 \pm 0.08$                       |
| Muons        |                        |                        |                                       |                            |                        |                                       |
|              | $\Delta R(l, j) > 0.4$ | $\parallel_{p_T > 40}$ | $\parallel_{p_T^{jet}/p_T^\mu < 0.5}$ | $\Delta R(l, j) > 0.4$     | $\parallel_{p_T > 40}$ | $\parallel_{p_T^{jet}/p_T^\mu < 0.5}$ |
| SU4          | $64.0 \pm 0.5$         | $75.5 \pm 0.6$         | $76.4 \pm 0.6$                        | $0.23 \pm 0.02$            | $0.42 \pm 0.02$        | $0.31 \pm 0.02$                       |
| $t\bar{t}$   | $70.3 \pm 0.4$         | $82.3 \pm 0.5$         | $83.8 \pm 0.5$                        | $0.42 \pm 0.03$            | $0.64 \pm 0.03$        | $0.49 \pm 0.03$                       |

Table 9.5: Efficiencies and fake rates for combinations of lepton jet separation  $\Delta R(l, j) > 0.4$  and additional kinematical cuts using Anti- $K_T$  jets. The  $\parallel$  sign means logical OR combination of the lepton jet separation requirement and  $p_T > 40$  and  $p_T^{jet}/p_T^\mu < 0.5$

## 9.4 Significance

This analysis will use the significances introduced in Section 8.1.1. The simple  $S_0$  given by Eq. 8.1 (page 161) is a fairly good approximation in situations when both  $S$  and  $B$  are significantly larger than unity.

This is, however, not always the case and we introduce here a different expression which starts from assuming that the number of observed events follows the Poisson distribution. From this it is possible to construct a log-likelihood ratio (LLR). In order to know the significance of the measurement one can use the  $\chi^2$  distribution and the fact that  $1/2$  of the log-likelihood corresponds to one unit of  $\Delta\chi^2$ . The resulting expression for the significance is

$$Z_{LLR} = \sqrt{2LLR} = \sqrt{2[(S+B)\ln(1+\frac{S}{B}) - S]}. \quad (9.7)$$

Contrary to  $S_0$  it is also valid when  $B$  is small compared to  $S$ . In the limit when  $S \ll B$  and  $S$  and  $B$  are known precisely, it is well approximated by  $S_0$  in Eq. 8.1 [174].

However, none of these two significance measures take into account the background uncertainty which is expected to be non-negligible, especially during the analysis of early data. Section 8.1.1 introduced the simple  $S_N$  in Eq. 8.2

A measure which is considered to include the background uncertainty in a more correct way is the  $Z_N$  given by Eq. 8.3 (Section 8.1.1). For the early 10 TeV analysis it is common to assume a background uncertainty of 100% or 50%. Although it may be a pessimistic choice, an uncertainty of 100% is chosen as default for this analysis. While the other significance measures may have limited validity,  $Z_N$  is a good measure for a wide range of situations.

Figure 9.6 shows a comparison of the presented measures of significance. In Figure (a) the x-axis represents the number of background events while the y-axis the number of signal events required for a  $5\sigma$  discovery. The comparison shows a clear discrepancy between the measures that take background uncertainty into account and those that do not. The simple expressions  $S_N$  is a fairly good, although optimistic approximations of the more complex  $Z_N$ . Looking at the measures without background uncertainty,  $S_0$  requires only slightly lower number of signal events for  $5\sigma$  than  $Z_{LLR}$

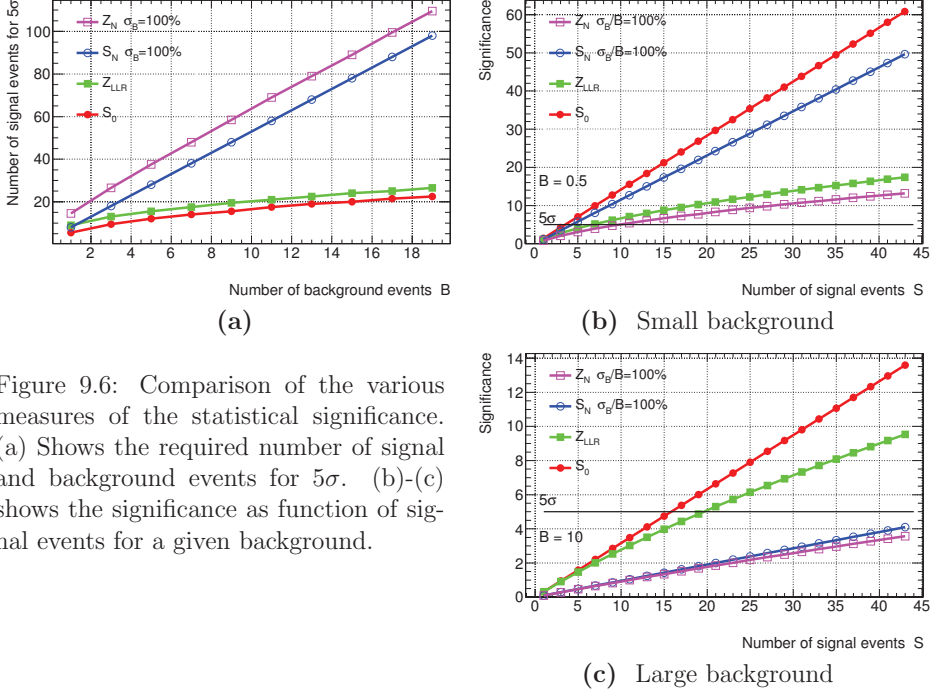


Figure 9.6: Comparison of the various measures of the statistical significance. (a) Shows the required number of signal and background events for  $5\sigma$ . (b)-(c) shows the significance as function of signal events for a given background.

Figure 9.6 (b) and (c) shows the significance as a function of number of signal events for a given number of background events. For small background (b) the effect of background uncertainty is also small and we observe that  $S_0$  and  $S_N$  are close, while  $Z_{LLR}$  gives approximately the same values as  $Z_N$ .

For high number of background events (c), the background uncertainty becomes a more important factor and reduces the significances given by expressions which include it,  $S_N$  and  $Z_N$ . In such case, the simple  $S_N$  is a fairly good approximation of  $Z_N$ . One observes also that with larger background,  $Z_{LLR}$  is closer to  $S_0$  which is an approximation of  $Z_{LLR}$  in the limit  $S \ll B$ .

As a part of the study of systematic effects both the cut optimization and the presentation of the analysis results use all four measures.

## 9.5 Jet inclusive trilepton searches

The jet inclusive trilepton search at 10 TeV is aimed at SU4, the low mass benchmark scenario, where the three leptons are mainly produced in long cascade decays accompanied by jets and  $\cancel{E}_T$ .

Figures 9.7 shows the  $p_T$  distributions for the three leading leptons after requiring three or more leptons in the event and  $M_{l+l-} > 20$  GeV. The total SM background includes contributions from all background samples while only the most prominent are showed explicitly. Compared to Figure 8.2 where dibosons and  $Zb$  were the dominant backgrounds, now the leptonically decaying  $Z$  accompanied by jets is the largest.

Figure 9.8 (a) shows the  $\cancel{E}_T$  after requiring  $N_l \geq 3$  and  $M_{l+l-} > 20$  GeV. The SUSY

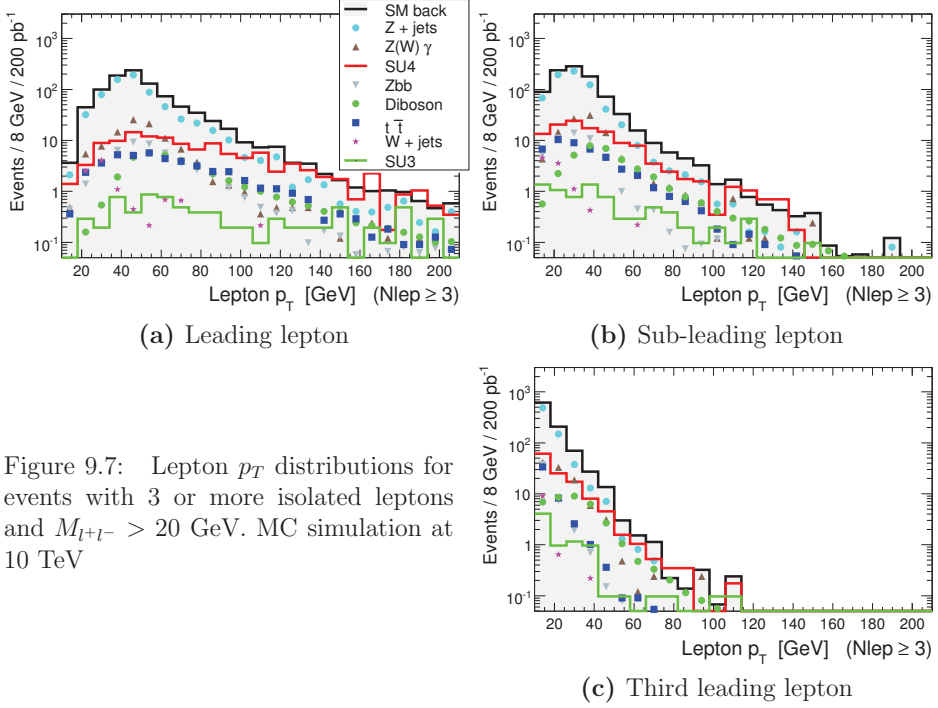


Figure 9.7: Lepton  $p_T$  distributions for events with 3 or more isolated leptons and  $M_{l+l-} > 20$  GeV. MC simulation at 10 TeV

signal, especially benchmark points dominated by long decay chains, like in this case SU4 and SU3, have a characteristically high  $\cancel{E}_T$  due to the escaping LSPs. Mass differences between the particles in the decay chain play also an important role. Among the SM backgrounds, it is only the  $t\bar{t}$  and diboson ( $WW$ ,  $WZ$  and  $ZZ$ ) signals that can be said to have significant  $\cancel{E}_T$ , basically due to escaping neutrinos. The potentially threatening backgrounds like  $Z$ +jets,  $W$ +jets,  $Zbb$  and  $Z(W)$  plus photon have much smaller  $\cancel{E}_T$ , typically below 50 GeV. While the analysis based on the limited set of 14 TeV MC samples indicated that hard cut on the leading jet basically eliminated the SM background, the extended set of backgrounds shows that a  $\cancel{E}_T$  cut is very efficient in rejecting a broad range of SM processes. The importance of this cut in searches for trilepton signal at LHC has also been pointed out in [169].

The other characteristic of many SUSY scenarios is a high number of jets. Figure 9.8 (b) shows the number of jets ( $p_T^{jet} > 10$  GeV, Anti- $K_T$  jets) for the two benchmark points SU4 and SU3 and various backgrounds, while Figure 9.9 (a-b) the  $p_T$  of the two leading jets. Although the SU3 sample is generated with low statistics and has a lower cross section at 10 TeV, it is still possible to recognize the very hard jet  $p_T$  spectrum which was exploited in the 14 TeV event selection. However, with the new set of SM backgrounds and focus on the low mass point SU4 rather than SU3, it has been necessary to revise the event selection.

The event selection cuts for the jet inclusive search for SU4 have been optimized through a three dimensional scan varying the cut on the first two leading jets and  $\cancel{E}_T$  in steps of 10 GeV up to 200 GeV for each of the variables. The cut optimization scan has been studied for:

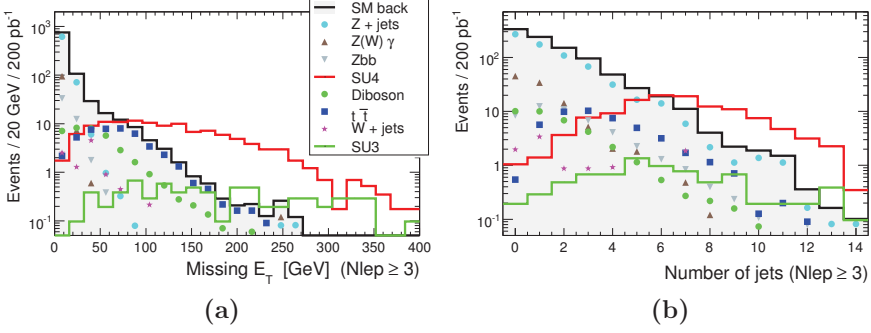


Figure 9.8:  $\cancel{E}_T$  distributions (a) and number of jets ( $p_T^{jet} > 10$  GeV, Anti- $K_T$ ) (b) after requiring  $N_l \geq 3$  and  $M_{l+l-} > 20$  GeV.

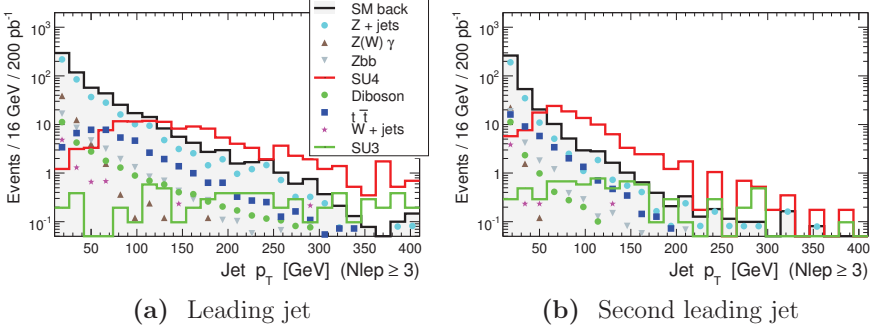


Figure 9.9:  $p_T$  distributions for the leading jet (a) and second leading jet (b) after requiring  $N_l \geq 3$  and  $M_{l+l-} > 20$  GeV.

- Cone and Anti- $K_T$  jets
- For the three requirements on leptons within  $\Delta R < 0.4$  to a jet in Equations 9.2-9.4.
- For all four significances, Eq. 8.1-8.3
- Background uncertainty  $\sigma_B/B$ : 100% and 50%

The full listing of the results of this study is shown in Chapter 12, Table 12.1.

Optimizing for the significance measures which do not take into account the background uncertainty,  $S_0$  from Eq. 8.1, and  $Z_{LLR}$  from Eq. 9.7, suggests milder cuts which give a relatively high number of both signal and background events. The  $S_N$  given by Eq. 8.2 on the other hand, favours harsh cuts which maximally reduce the background, especially when assuming 100% background uncertainty.

The optimization for  $Z_N$  (Eq. 8.3 with 100% background uncertainty) falls between these two extremes and reaches often a maximum close to the cuts suggested by the optimization for  $S_0$  or  $S_N$  with 50% background uncertainty. The  $Z_N$  significance is the most stable with respect to the varying lepton-jets separation requirements and jet algorithms. As it in addition gives moderate cut values which are also favourable for the other significance measures, it is chosen as the basis for the event selection optimization. Moderate



cuts of jet  $p_T$  and  $\cancel{E}_T$  are also considered more reliable due to the large uncertainties, especially for high values, which are expected in the early analyses.

The background uncertainty was set to 100%, however 50% was also studied. The  $Z_N$  significance is not very sensitive to this change and gives similar cut values for both uncertainties. With  $\sigma_B/B = 100\%$ ,  $S_N$  obtains highest values with minimal background often obtained with very harsh cuts which make the result sensitive to statistical fluctuations and uncertainties related to the tails of distributions. In such cases the chosen event selection is the one giving highest significance with cuts not higher than about 100 GeV. Using  $\sigma_B/B = 50\%$   $S_N$ , tends also towards milder cuts.

The optimization scans have been performed for the three lepton-jet separation requirements described in Section 9.3.1 as well as for both the fixed  $\Delta R = 0.4$  cone and Anti- $K_T$  jets. Comparing the maximum significances obtained in the scan, the simple rejection of leptons closer to a jet than  $\Delta R = 0.4$ , especially when using the Anti- $K_T$  jets, gives the lowest values. The combined cut in Eq. 9.3 improves the significance, while the highest maximum significances are obtained with cut in Eq. 9.4. These observations are valid both for the cone and Anti- $K_T$  jets, while the latter gives slightly better results.

Based on the efficiency and fake rate studies as well as these cut scans, the cut in Eq. 9.4 has been selected as the base-line muon-jet separation requirement. Since only a very small fraction of the electrons fall within  $0.2 < \Delta R(e, j) < 0.4$  it is not applied on electrons.

The Anti- $K_T$  jets show the best results and are therefore chosen as the default jet algorithm in the final analysis. As shown in Figure 9.3, Anti- $K_T$  jets are softer than the cone jets and the optimization suggests cuts which are 10-20 GeV lower than for the cone jets.

The resulting event selection optimised for SU4 is:

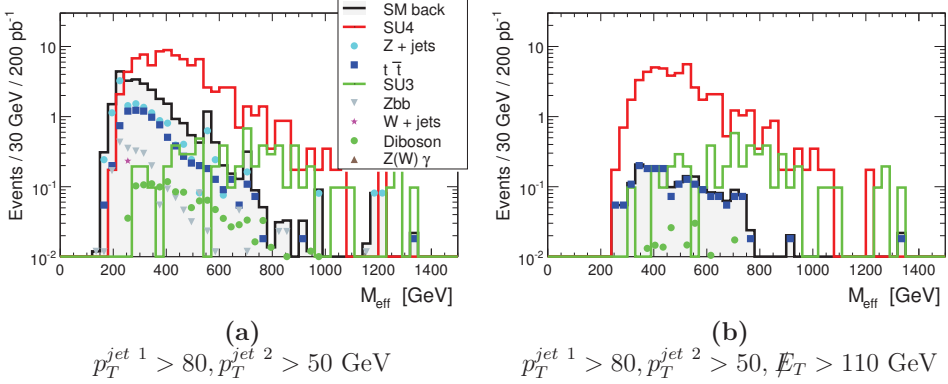
- $p_T^{jet\ 1} > 80\text{ GeV}$ ,  $p_T^{jet\ 2} > 50\text{ GeV}$  (Anti- $K_T$ )
- $\cancel{E}_T > 110\text{ GeV}$

Optimized cuts suggested by the same scan for the SU3 benchmark point are not reliable due to the very low number of simulated events. However, due to the characteristic phenomenology of the point, the scan favours much harder cuts than for SU4, well above 100 GeV both on the  $p_T$  of the two leading jets and  $\cancel{E}_T$ . This is consistent with the results from the 14 TeV analysis.

Table 9.6 shows the cut flow for the various signal samples and backgrounds normalized to  $200\text{ pb}^{-1}$  which is defined by ATLAS as an early 10 TeV scenario. After the jet- $p_T$  cuts the background from  $Z$ +jets processes is at the level of the  $t\bar{t}$ , while it is completely removed by the  $\cancel{E}_T$  cut. The final background is only due to  $t\bar{t}$  and diboson events. The latter is combined in order to simplify the table, however it is the  $WZ$  production which is the main contributor.

The rightmost columns list the statistical significances as given by the formulas introduced in Section 9.4. With  $200\text{ pb}^{-1}$  of data at 10 TeV it is only the low mass SUSY scenario represented by SU4 which may reach a  $5\sigma$  significance.

The 100% background uncertainty is a conservative choice and the table shows also significances for 50%. For SU4 the  $Z_N$  from Eq. 8.3 increases with 23% while the other points only about 4%. For all points,  $S_N$  from Eq. 8.2 increases by 45%. The improvement demonstrates that good understanding of background will be a very important.

Figure 9.10:  $M_{eff}$  after the successive event selection.

For SU3 which has the second highest cross section the selected cuts are too soft. However, even with optimized cuts the most conservative significance  $Z_N$  is only around 2. The other significance measures may obtain values close to 5, but the required hard cuts and the low MC statistics make the results very sensitive to statistical fluctuations.

The effective mass is a powerful discriminating quantity capable of showing a SUSY excess over the SM background. As this analysis is dealing with two jets it is defined as:

$$M_{eff} = \cancel{E}_T + \sum_{i=1,2} p_T^{jet\ i} + \sum_{i=1}^{N_{lep}} p_T^{lep\ i} \quad (9.8)$$

where the sum over jet  $p_T$  runs only over the two leading ones. Figure 9.10 shows the  $M_{eff}$  distributions after the successive event selection cuts. After the cut on the  $p_T$  of the two leading jet, Figure 9.10 (a), the background from the  $Z$ +jets samples is comparable to  $t\bar{t}$  and there is a smaller contribution from a number of other processes. Figure 9.10 (b) shows the same distribution after applying all event selection cuts. We observe that the  $\cancel{E}_T$  cut removes efficiently all but  $t\bar{t}$  and diboson backgrounds. Both SU4 and SU3 signals are visible above the background.

### 9.5.1 Variation of the jet inclusive analysis

The analysis was then repeated changing various parameters.

#### Cone jets

The cut flow and results for the analysis using cone jets is shown in Table 9.7. The significances are slightly lower as compared to the analysis with  $Anti-K_T$  jets. Even with cuts optimized for cone jets which implies a slightly higher cut on the second leading jet  $p_T^{jet\ 2} > 60$  GeV, the significances are very close to those in Table 9.6. With this cut the significance  $Z_N$  for SU4 is 9.2.

#### NLO - LO cross sections

In order to see the effect of using NLO cross section, the significances were calculated without using K-factors. As the main backgrounds,  $t\bar{t}$  and  $WZ$ -diboson events, are generated with an NLO generator, the cross sections were divided by the respective K-factors,  $WZ$ :

|                 | $N_l \geq 3$ | Jet 1,2 $p_T$ | $\cancel{E}_T > 110$ | $S_0$              | $Z_{LLR}$ | $S_N$ | $Z_N$ |
|-----------------|--------------|---------------|----------------------|--------------------|-----------|-------|-------|
|                 |              |               |                      | $\sigma_B = 100\%$ |           |       |       |
| SU4             | 131.1        | 88.6          | 50.0                 | 32.8               | 16.6      | 18.0  | 9.2   |
| SU3             | 8.1          | 6.4           | 5.0                  | 3.3                | 3.5       | 1.8   | 1.5   |
| SU1             | 4.3          | 3.2           | 3.0                  | 1.9                | 2.4       | 1.1   | 0.9   |
|                 |              |               |                      | $\sigma_B = 50\%$  |           |       |       |
| SU4             | 131.1        | 88.6          | 50.0                 |                    |           | 26.1  | 11.4  |
| SU3             | 8.1          | 6.4           | 5.0                  |                    |           | 2.6   | 2.0   |
| SU1             | 4.3          | 3.2           | 3.0                  |                    |           | 1.5   | 1.2   |
| $Zbb$           | 50.8         | 2.3           | 0.0                  |                    |           |       |       |
| $W + jet$       | 10.1         | 0.2           | 0.0                  |                    |           |       |       |
| $Z, W + \gamma$ | 105.9        | 0.0           | 0.0                  |                    |           |       |       |
| Single top      | 4.6          | 0.4           | 0.0                  |                    |           |       |       |
| $Z + jets$      | 734.5        | 14.3          | 0.0                  |                    |           |       |       |
| $WW, WZ, ZZ$    | 36.3         | 1.5           | 0.3                  |                    |           |       |       |
| $Wbb$           | 0.2          | 0.1           | 0.0                  |                    |           |       |       |
| $t\bar{t}$      | 48.8         | 9.7           | 2.0                  |                    |           |       |       |
| QCD $b\bar{b}$  | 0.0          | 0.0           | 0.0                  |                    |           |       |       |
| Sum B           | 991.2        | 28.6          | 2.3                  |                    |           |       |       |

Table 9.6: Cut flow jet inclusive search at 10 TeV and significance for the various signal samples. All numbers are normalized to 200 pb<sup>-1</sup>. Anti- $K_T$  tower jets.

1.5 [175] and  $t\bar{t}$ : 2.05 [78]. Table 9.8 summarizes the significances for LO cross sections. The K-factors for the dominant backgrounds are higher than for signal samples. This is mirrored in  $S_N$  which is higher for LO than for NLO (approximately 7%). The remaining significances are roughly 10% lower for LO.

### Lepton isolation

The applied commonly used calorimeter-based isolation cut is mild and not optimised for a multilepton analysis. Applying instead the well performing relative track-based isolation which was discussed in Chapter 7 with a cut  $p_{T\Sigma\text{track}}^{\Delta R=0.3}/p_T < 0.11$  reduces the  $WZ$  background by 34% and  $t\bar{t}$  by 85%. The signal is also reduced, in particular SU4 which is reduced by 50%. The resulting significances are listed in Table 9.9. The  $Z_{LLR}$  significance measure which favours high numbers of both signal and background events is reduced. On the other hand, the other significance measures clearly increase. For SU4,  $S_N$  and  $S_0$  increase by 70% and 13%, respectively. This shows again the importance of the optimization of isolation cuts as one of the main tools in reduction of the large  $t\bar{t}$  background. However, this comparison is not entirely fair. Therefore Table 9.10 shows the corresponding numbers using a more optimized calorimeter-based isolation:  $E_{\text{cal}}^{\Delta R=0.2} < 5$  GeV. The choice of the cut is suggested by the optimization study in Section 7.4. Compared to the track-based isolation, all significance measures except  $Z_N$ , are lower, in most cases of order of 10%. However, compared to the default cut,  $E_{\text{cal}}^{\Delta R=0.2} < 10$  GeV (Table 9.6) for SU4  $S_0$  and  $Z_N$  increase by about 5%, while  $S_N$  which is enhanced by the reduced background increases by 20%. The remaining points follow the same pattern. The  $Z_N$  measure for SU3 and SU1 gains as much as 30% and 60%, respectively.

|       | $N_l \geq 3$ | Jet 1,2 $p_T$ | $\cancel{E}_T > 110$ | $S_0$ | $Z_{LLR}$ | $S_N$ | $Z_N$ |
|-------|--------------|---------------|----------------------|-------|-----------|-------|-------|
| SU4   | 137.7        | 94.1          | 52.0                 | 32.4  | 16.8      | 17.1  | 9.1   |
| SU3   | 8.5          | 6.6           | 5.2                  | 3.3   | 3.5       | 1.7   | 1.4   |
| SU1   | 4.3          | 3.2           | 3.0                  | 1.9   | 2.4       | 1.0   | 0.9   |
| Sum B | 1250.4       | 29.6          | 2.6                  |       |           |       |       |

Table 9.7: Same as Table 9.6, but using cone jets.

|       | $N_l \geq 3$ | Jet 1,2 $p_T$ | $\cancel{E}_T > 110$ | $S_0$ | $Z_{LLR}$ | $S_N$ | $Z_N$ |
|-------|--------------|---------------|----------------------|-------|-----------|-------|-------|
| SU4   | 95.7         | 64.6          | 36.5                 | 29.8  | 14.5      | 18.9  | 8.7   |
| SU3   | 6.1          | 4.8           | 3.8                  | 3.1   | 3.1       | 2.0   | 1.5   |
| SU1   | 3.2          | 2.4           | 2.2                  | 1.8   | 2.1       | 1.2   | 0.9   |
| Sum B | 787.6        | 21.4          | 1.5                  |       |           |       |       |

Table 9.8: Same as Table 9.6, but using the LO cross sections.

|       | $N_l \geq 3$ | Jet 1,2 $p_T$ | $\cancel{E}_T > 110$ | $S_0$ | $Z_{LLR}$ | $S_N$ | $Z_N$ |
|-------|--------------|---------------|----------------------|-------|-----------|-------|-------|
| SU4   | 63.8         | 44.5          | 25.5                 | 37.2  | 13.5      | 30.7  | 9.6   |
| SU3   | 4.7          | 3.7           | 2.8                  | 4.1   | 3.2       | 3.4   | 2.0   |
| SU1   | 2.8          | 2.0           | 1.9                  | 2.8   | 2.4       | 2.3   | 1.5   |
| Sum B | 351.5        | 10.1          | 0.5                  |       |           |       |       |

 Table 9.9: Same as Table 9.6, but using the track-based isolation  $p_{T\Sigma\text{track}}^{\Delta R=0.3}/p_T < 0.11$ .

|       | $N_l \geq 3$ | Jet 1,2 $p_T$ | $\cancel{E}_T > 110$ | $S_0$ | $Z_{LLR}$ | $S_N$ | $Z_N$ |
|-------|--------------|---------------|----------------------|-------|-----------|-------|-------|
| SU4   | 113.8        | 76.5          | 43.4                 | 34.7  | 16.1      | 21.6  | 9.7   |
| SU3   | 7.0          | 5.4           | 4.5                  | 3.6   | 3.4       | 2.2   | 2.0   |
| SU1   | 3.7          | 2.8           | 2.7                  | 2.2   | 2.4       | 1.4   | 1.1   |
| Sum B | 829.0        | 20.9          | 1.6                  |       |           |       |       |

 Table 9.10: Same as Table 9.6, but using tighter calorimeter-based isolation  $E_{\text{cal}}^{\Delta R=0.2} < 5$  GeV.

|       | $N_l \geq 3$ | Jet 1,2 $p_T$ | $\cancel{E}_T > 110$ | $S_0$ | $Z_{LLR}$ | $S_N$ | $Z_N$ |
|-------|--------------|---------------|----------------------|-------|-----------|-------|-------|
| SU4   | 131.1        | 88.6          | 50.0                 | 23.2  | 14.8      | 9.8   | 6.7   |
| SU3   | 8.1          | 6.4           | 5.0                  | 2.3   | 3.0       | 1.0   | 0.9   |
| SU1   | 4.2          | 3.2           | 3.0                  | 1.4   | 2.1       | 0.6   | 0.5   |
| Sum B | 1982.4       | 57.1          | 4.7                  |       |           |       |       |

Table 9.11: Same as Table 9.6, but SM background is multiplied by a factor of 2.

|           | SU4 (100%) | SU4 (50%) | SU3 (20%) | SU1 (20%) |
|-----------|------------|-----------|-----------|-----------|
| $Z_N$     | 38         | 28        | 1200      | 7000      |
| $S_N$     | 5          | 5         | 600       | 3500      |
| $Z_{LLR}$ | 23         |           | 750       | 2000      |
| $S_0$     | 5          |           | 460       | 1400      |

Table 9.12: Integrated luminosity required for  $5\sigma$  discovery. Luminosity is given in  $\text{pb}^{-1}$ . The numbers in parenthesis give the background uncertainty.

### SM background

Knowing that the SM background will be difficult to estimate during the early analysis, Table 9.11 shows the significances with SM increased by a factor two. SU4 reaches still  $5\sigma$ , however both  $Z_N$  and  $S_0$  are deduced by 30%.  $Z_{LLR}$  is least affected being reduced by only 10% while  $S_N$  is reduced by 45%.

## 9.6 Conclusions

Given the original start-up scenario of LHC operation, with a planned center of mass energy of 10 TeV during the early data taking, it has been necessary to revise the jet inclusive trilepton search. From the MC point of view, the new simulation offered a number of samples as the earlier applied filter cuts which were incompatible with the trilepton event selection were removed. Despite the new situation, this analysis shows that at least a low mass SUSY scenario like SU4 has a promising significance of  $Z_N = 9.2$  with  $200 \text{ pb}^{-1}$  of integrated luminosity. Based on the number of signal and background events in Table 9.6, extrapolation has been made in order to find the required integrated luminosity for  $5\sigma$  discovery. With 100% background uncertainty, SU4 reaches  $5\sigma$  ( $Z_N$ ) after  $38 \text{ pb}^{-1}$ , while with 50% already after  $28 \text{ pb}^{-1}$ . Table 9.12 summarizes the required luminosities for all points and the various significance measures.

Points SU3 and SU1 have a relatively low  $S/B$  ratio and with 100% and 50% background uncertainty this becomes the dominant factor. For SU3,  $Z_N$  and  $S_N$  flattens out and reaches about 2 and 4 for  $\sigma_B/B = 100\%$  and  $\sigma_B/B = 50\%$ , respectively. The corresponding approximate numbers for SU1 are 1 and 3. In the 14 TeV analysis the background uncertainty was estimated to 20% at  $1 \text{ fb}^{-1}$  [78]. This number was used in the estimates quoted in Table 9.12. With this lower background uncertainty, SU3 and SU1 can reach a  $5\sigma$  discovery potential with integrated luminosity of several  $\text{fb}^{-1}$ . However, this analysis is not optimal for these points as these have harder jet and higher  $\cancel{E}_T$ .

With the proposed object definition, especially in the field of muon jet separation, the results of the analysis are very similar when using the traditional fixed  $\Delta R = 0.4$  cone and the Anti- $K_T$  jets. The discovery potential improves significantly with more optimized isolation requirements. Here the analysis tested both a track- and calorimeter-based cut.

In order to estimate the impact of the background uncertainty on the significances given by  $Z_{LLR}$  (Eq. 9.7) and  $S_N$  (Eq. 8.2), the SM background was multiplied by two. SU4 reaches still  $5\sigma$  within  $200 \text{ pb}^{-1}$ , but the significances are in general strongly reduced.



# Chapter 10

## MSSM24 multilepton analysis with early data at 10 TeV

So far most of the SUSY searches performed in ATLAS have focused on the constrained mSUGRA models. By choosing a mechanism for SUSY breaking and assuming unification at GUT scale one can reduce the number of model parameters from the original 105 for MSSM to 4-5. While this is a clear advantage from the simulation point of view, it also imposes some crucial phenomenological constraints. Section 3.6.2 showed how the GUT-scale unification imposes the very characteristic mass hierarchy of mSUGRA:  $m_{\tilde{g}} : m_{\tilde{\chi}_2^0} : m_{\tilde{\chi}_1^0} = 7 : 2 : 1$ . Figure 6.9, Case (i) showed a typical mSUGRA sparticle composition and mass hierarchy, which influences how left-/right-handed squarks decay and makes the  $\chi_{3,4}^0$  and  $\chi_2^\pm$  inaccessible in most squark decays. As a consequence, most standard SUSY searches in ATLAS follow the same pattern exploiting the high number of hard jets and considerable  $\cancel{E}_T$  expected from a decay chain originating from gluino or squark production.

Although the various benchmark points cover a broad spectrum of SUSY scenarios, Nature itself may be much more complex. The full parameter space of MSSM is overwhelming, however many of the 105 parameters cause CP violation or flavor-changing neutral current processes and are thus tightly constrained [176]. If one keeps only processes that conserve CP and flavor, one is left with 24 parameters: gaugino and Higgsino masses ( $M_1, M_2, M_3, \mu$ ), slepton and squark masses, Higgs potential terms ( $\tan \beta, m_A$ ), and trilinear terms ( $A_\tau, A_t, A_b$ ).

ISASUSY [139] provides a tool which allows to vary these 24 MSSM parameters providing thus the possibilities to break some of the limitations of mSUGRA and study different phenomenologies.

This framework has been used to construct a number of new benchmark points ordered in a phenomenological grid (PhenoGrid) [177]. These points try in particular to challenge the common mSUGRA patterns. As a first step it is interesting to study how well the standard analyses developed for mSUGRA perform when facing more complex and in many cases different scenarios. As a next step it will be necessary to develop new strategies optimized for the new picture. However, the present feasibility study will only deal with the first question.

## 10.1 Gaugino production and trilepton signal in MSSM24

The PhenoGrid is a simulation of a new set of SUSY benchmark points which covers a wide range of scenarios. However, those which are most relevant for the trilepton analysis are mostly related to gaugino production. In the MSSM24 framework the gaugino masses and composition are defined by  $M_1$ ,  $M_2$ ,  $\mu$  and  $\tan\beta$ . Although the aim is to break the mSUGRA hierarchy, the relation  $M_1 < M_2 < \mu$  is kept, such that the LSP in most cases is still pure bino like. The usual mass degeneracy of selectrons and smuons is not enforced. Below follows a description of the phenomenology classes most suitable for trilepton analysis.

Direct gaugino production becomes important in scenarios where the gluino and squarks are heavy. The only mSUGRA benchmark point with this feature was SU2 studied in the 14 TeV analysis, where the gluino is relatively accessible ( $m_{\tilde{g}} = 856.59$  GeV). However, in PhenoGrid gaugino points the gluino is also beyond the energy reach. Table 10.1 shows an overview of some of the sparticle masses for the various points. A full listing of the cross sections can be found in Appendix C Table C.3. Three direct gaugino classes have been defined:

- The sleptons are also heavy and the gauginos decay to leptons via real or virtual  $W$ ,  $Z$  or  $h$ , depending on the phase space (label: directgaugino, 4 points)
- In the two other classes the sleptons or the staus are lighter than some of the gauginos which thus preferably decay via the available slepton creating a small cascade (label: directgaugino\_sle, directgaugino\_stau, 12 points in each class)

| Sample             | $\sigma_{LO}$ [pb] | $m_{\tilde{\chi}_1^0}$ [GeV] | $m_{\tilde{\chi}_2^0}$ | $m_{\tilde{\chi}_1^\pm}$ | $m_{\tilde{\tau}_1}$ | $m_{\tilde{\mu}/\tilde{e}}$ |
|--------------------|--------------------|------------------------------|------------------------|--------------------------|----------------------|-----------------------------|
| directgaugino      | 1.9 - 19.0         | 37 - 60                      | 157 - 181              | 157 - 181                | 487 - 573            | $\sim 530$                  |
| directgaugino_sle  | 0.5 - 8.2          | 56 - 100                     | 160 - 300              | 160 - 300                | 150 - 366            | 127 - 150                   |
| directgaugino_stau | 0.3 - 9.0          | 60 - 100                     | 120 - 300              | 120 - 300                | 110 - 285            | $\sim 530$                  |

Table 10.1: Overview of the MSSM24 points with direct gaugino production. The masses of the gluino and squarks are heavier than 4500 GeV for all points, and inaccessible. The mass of  $\tilde{\chi}_1^\pm$  and  $\tilde{\chi}_2^0$  are equal for all points in each of the three classes. [177]

K-factors were not available for the MSSM24 points. However, as one can expect them to be larger than one, using LO cross sections for signal and NLO for SM backgrounds leads to underestimated results.

The combination of masses is always such that the desired decay process is kinematically allowed. In the first category of direct-gaugino points (no sleptons available in the decay) the dominant pair production processes are  $\tilde{\chi}_2^0\tilde{\chi}_1^\pm$  with approximately 65% and  $\tilde{\chi}_1^\pm\tilde{\chi}_1^\mp$  35%. In the two other categories of points, direct-gaugino with decay via sleptons and direct-gaugino with decay via staus, these two account for a slightly lower rate of approximately about 60% and 30% as direct slepton and stau production is also possible.

In the two first direct-gaugino points,  $\tilde{\chi}_1^\pm$  and  $\tilde{\chi}_2^0$  decay mainly to quarks and a lighter gaugino, usually the LSP. In the third point,  $\tilde{\chi}_1^\pm$  and  $\tilde{\chi}_2^0$  are significantly heavier than  $W$  and  $Z$  ( $m_{\tilde{\chi}} = 156$  GeV) and decay via these. While in the fourth point where they are



even heavier ( $m_{\tilde{\chi}} = 181$  GeV), gauginos may also decay via Higgs ( $\tilde{\chi}_1^\pm \rightarrow W^\pm \tilde{\chi}_1^0(100\%)$ ,  $\tilde{\chi}_2^0 \rightarrow h\tilde{\chi}_1^0(82\%), Z\tilde{\chi}_1^0(17\%)$ ).

All categories of direct-gaugino points are characterized by a considerable cross section for multilepton final states with relatively few and soft jets, as well as small  $\cancel{E}_T$ . Figure 10.1 (a) shows the number of leptons in these points and (b) the lepton  $p_T$  after requiring  $N_l \geq 3$  and  $m_{l+l-} > 20$  GeV. Figure 10.2 shows the same distributions for the most promising direct-gaugino-slepton points (gaugino decay via sleptons). Figure (a) shows that there is a significant fraction of events with three or more leptons. Plot (b) shows the lepton  $p_T$  distributions (SM background not included as in Figure 10.1 (b)). We observe that there are two points that have characteristically softer leptons which can be explained by the small mass differences between the  $\tilde{\chi}_2^0$  and the  $\tilde{e}_R, \tilde{\mu}_R$  via which it decays, and  $\tilde{\chi}_1^0$ .

The characteristics of the phenomenology represented by the direct-gaugino families make them interesting in the context of jet exclusive trilepton analysis.

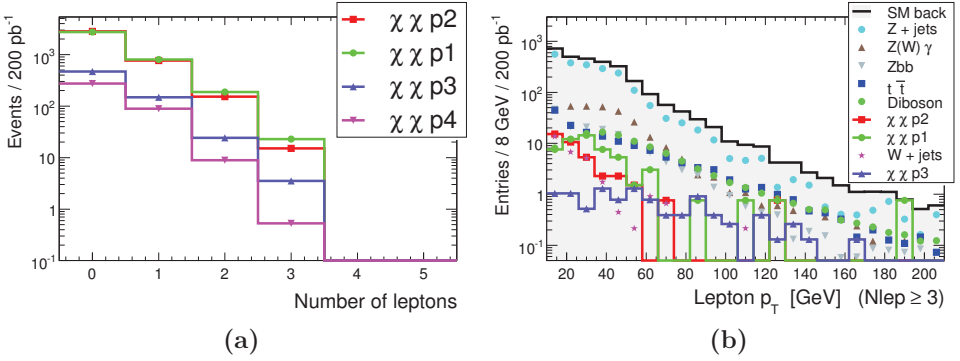


Figure 10.1: Number of leptons (a) and lepton  $p_T$  (b) after requiring  $N_l \geq 3$  and  $M_{l+l-} > 20$  GeV for direct-gaugino points (no decay via sleptons).

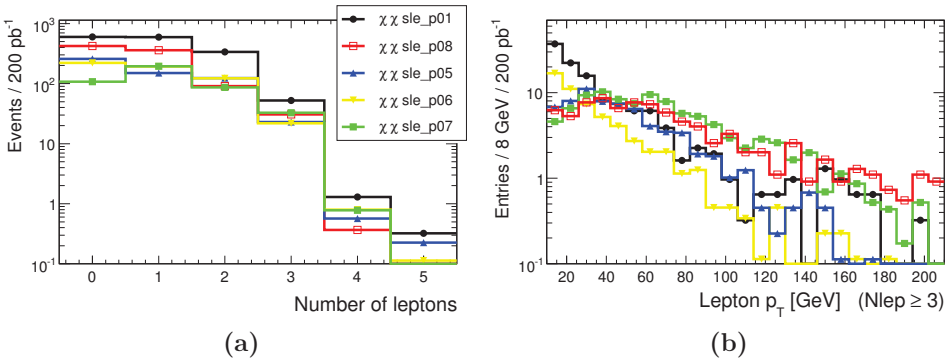


Figure 10.2: Number of lepton (a) and lepton  $p_T$  (b) after requiring  $N_l \geq 3$  and  $M_{l+l-} > 20$  GeV for direct-gaugino-slepton points (gauginos decay via sleptons). Only the most promising of the 12 sets are included. The SM background is similar to Figure 10.1 (b) and is omitted in order to make the plots more readable.

### 10.1.1 Jet exclusive analysis

The goal of this section is first to study how the jet exclusive search strategy developed earlier performs when applied on the various MSSM24 direct gaugino production scenarios.

Analysis objects are selected according to definitions from the previous section using the muon-jet separation requirement given by Eq. 9.4. The MSSM24 points are simulated using ATLFast II which does not provide the Anti- $K_T$  jets. The cone tower jets are therefore used.

In the baseline analysis developed for the CSC exercise the following cuts were applied [78]:

1. At least one pair of opposite sign, same flavor leptons ( $l^+l^-$ ) ( $e^+e^-$  or  $\mu^+\mu^-$ ) with  $M_{l^+l^-} > 20$  GeV
2.  $N_\ell \geq 3$  ( $\ell \in \{e, \mu\}$ )
3.  $p_{T\text{track,max}}^{\Delta R=0.2} < 2$  GeV for electrons,  $p_{T\text{track,max}}^{\Delta R=0.2} < 1$  GeV for muons, where  $p_{T\text{track,max}}^{\Delta R=0.2}(\ell)$  is the maximum  $p_T$  of any track in a  $\Delta R = 0.2$  cone around the lepton
4. No  $l^+l^-$  dilepton pair with invariant mass in the  $Z$ -mass window  $|M_{l^+l^-} - M_Z| > 10$  GeV
5.  $\cancel{E}_T > 30$  GeV – a moderate  $\cancel{E}_T$  cut
6. No jet with  $p_T > 20$  GeV – referred to as the Jet Veto

Cut 3, which is an additional isolation criterion, removes efficiently background due to *extra* leptons, typically from heavy flavor decay in processes like  $t\bar{t}$ . However, the cut reduces also the signal by about 50%. Section 10.1.2 will therefore present the results obtained with Cut 3 replaced by another track-based isolation variable.

Several of the SM backgrounds involve a  $Z$  and a cut on the invariant mass of a  $l^+l^-$  pair can be used to suppress it. In cases where it is possible to construct more than one  $l^+l^-$  pair, the cut is performed on the mass closest to the nominal  $Z$  mass.

Figure 10.3 shows variables used in the last event selection cuts. The  $\cancel{E}_T$  distribution after Cut 4 is shown in Figure 10.3 (a). For the direct gaugino scenarios it is moderate compared to more common mSUGRA scenarios and a minimum cut is set to 30 GeV. Cut 5 removes efficiently contributions from  $Z$ +jets,  $Z/W+\gamma$  and  $Zb\bar{b}$  while the  $t\bar{t}$  background is as expected almost unchanged.

The number of jets as well as the  $p_T$  of the leading jet are shown in Figure 10.3 (b) and (c) respectively. The distributions illustrate the low hadronic activity which is only found in the mSUGRA Focus point scenario (point SU2) making it fundamentally different from the common mSUGRA decay chains. As a last cut it is required that no jet in the event can have  $p_T^{\text{jet}} > 20$  GeV. Cut 6, the jet veto, rejects most of the  $t\bar{t}$  background leaving  $Z$ +jets as the largest background source. Normalized to 200 pb $^{-1}$   $Z$ +jets and  $t\bar{t}$  contribute with 0.91 and 0.20 events, respectively. This situation differs from the 14 TeV mSUGRA analysis where the  $Z$ +jets samples were not available and the background was dominated by diboson events, mainly  $WZ$ . In the current analysis the diboson contribution is similar to  $t\bar{t}$  (0.21 events normalized to 200 pb $^{-1}$ ).

Figure 10.4 shows the same distributions as Figure 10.3 for the most promising direct-gaugino-slepton points. Although one can observe more  $\cancel{E}_T$  as well and higher hadronic activity, the scenarios are still most relevant for the jet exclusive analysis.

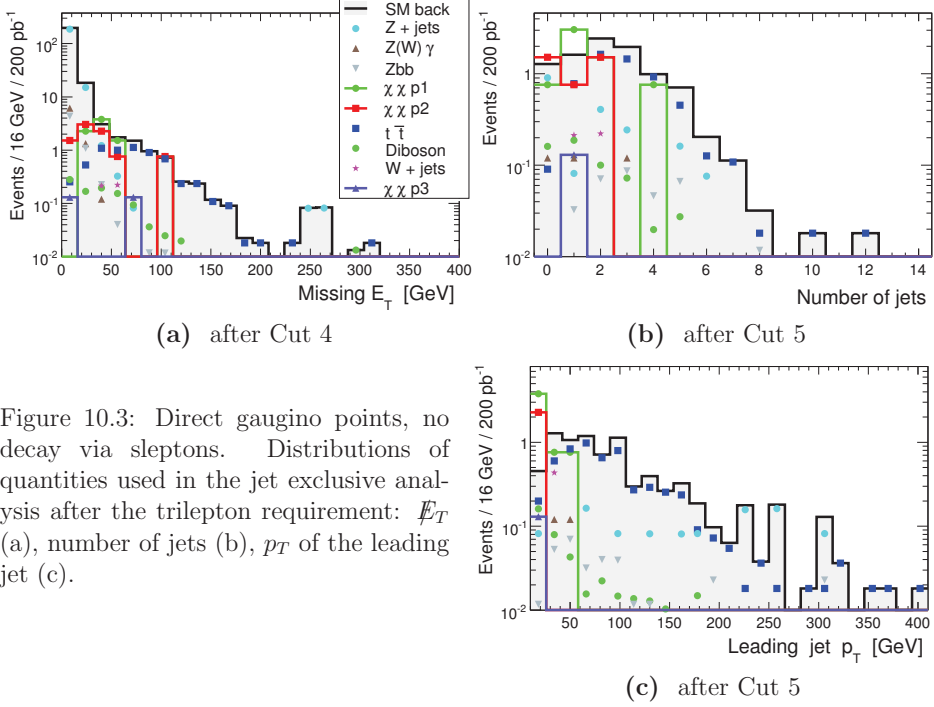


Figure 10.3: Direct gaugino points, no decay via sleptons. Distributions of quantities used in the jet exclusive analysis after the trilepton requirement:  $\cancel{E}_T$  (a), number of jets (b),  $p_T$  of the leading jet (c).

|                  | ev. after Cut 6 | $S_0$ | $Z_{LLR}$ | $S_N$ | $Z_N$ |
|------------------|-----------------|-------|-----------|-------|-------|
| directgaugino.p1 | 3.8             | 3.2   | 3.1       | 2.0   | 1.5   |
| directgaugino.p2 | 3.8             | 3.2   | 3.1       | 2.0   | 1.5   |
| directgaugino.p3 | 0.1             | 0.1   | 0.4       | 0.1   | 0.0   |
| directgaugino.p4 | 0.0             | 0.0   | 0.0       | 0.0   | 0.0   |
| Sum background   | 1.4             |       |           |       |       |

Table 10.2: Significances for the jet exclusive search at 10 TeV, direct-gaugino points. All numbers are normalized to 200 pb<sup>-1</sup>. 100% background uncertainty.

Contrary to the CSC analysis the jet veto is not optional in the present set of cuts. The current SM background composition and the significance measures, taking into account the background uncertainty ( $S_N$  and  $Z_N$ ) which favour low background level, it is clearly an advantage to apply the jet veto.

Table 10.2 summarizes the significances for the direct gaugino points without decay via sleptons. Despite the fact that points 1 and 2 have relatively high cross sections (19 pb) dominated by the production of pairs of lightest gauginos, there are only few trilepton events. The  $\tilde{\chi}_2^0$  and  $\tilde{\chi}_1^+$  are light and can only decay via virtual  $Z$  and  $W$  which then dominantly decay to quarks. Such scenarios will clearly require large data samples before one can claim any significant signal.

The significances of the direct gaugino points with decay via sleptons are summarized in Table 10.3. Particularly three points show good discovery potential within this analysis:

- Point 1:  $\sigma_{p1} = 8.1$  pb,  $m_{\tilde{\chi}_1^0} = 55$  GeV,  $m_{\tilde{\chi}_2^0, \tilde{\chi}_1^+} = 120$  GeV,  $m_{\tilde{e}_R} = 111$  GeV and

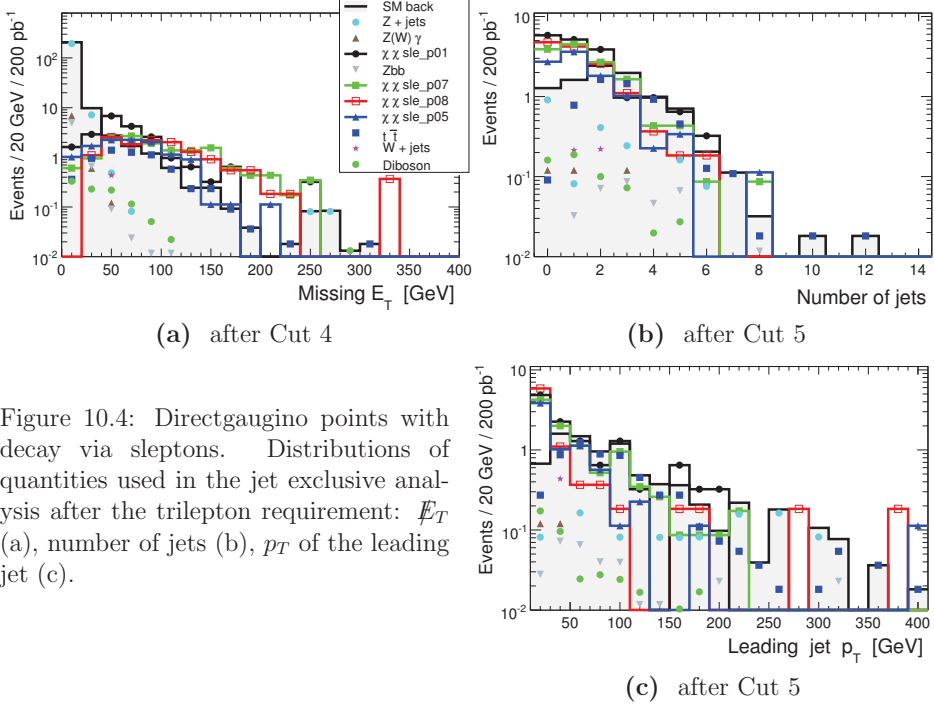


Figure 10.4: Directgaugino points with decay via sleptons. Distributions of quantities used in the jet exclusive analysis after the trilepton requirement:  $\cancel{E}_T$  (a), number of jets (b),  $p_T$  of the leading jet (c).

$m_{\tilde{\mu}_R} = 108$  GeV. The right-handed sleptons are much lighter than the left-handed ones ( $m_{\tilde{e}_L}, m_{\tilde{\mu}_L} \sim 440$  GeV), the  $\tilde{\chi}_2^0$  decays dominantly to  $\tilde{\mu}_R \mu$  and  $\tilde{e}_R e$  (63%, 33% respectively).  $\tilde{\chi}_1^+$  decays mainly to quarks, but also to the lightest  $\tilde{\tau}_1$  (20%) and to leptons,  $\tilde{\chi}_1^0 \nu_l$ , (9% per flavor  $e, \mu, \tau$ ).

- Point 7:  $\sigma_{p7} = 2.2$  pb,  $m_{\tilde{\chi}_1^0} = 62$  GeV,  $m_{\tilde{\chi}_2^0, \tilde{\chi}_1^+} = 200$  GeV,  $m_{\tilde{e}_L}, m_{\tilde{e}_R} \sim 133$  GeV  $m_{\tilde{\mu}_L}, m_{\tilde{\mu}_R} = 128$  GeV. Both left- and right-handed sleptons are light and only slightly heavier than the sneutrinos. The highest branching fractions are therefore to the  $\tilde{\nu}_l \nu_l$ , however approximately with the same rates,  $\tilde{\chi}_2^0$  decays to  $\tilde{\mu}_L \mu$  and  $\tilde{e}_L e$  (23%, 21% respectively).  $\tilde{\chi}_1^+$  decays also mainly to a sneutrino plus a lepton ( $\sim 30\%$  per flavor,  $e, \mu$ ). It also has a considerable branching fraction of approximately 16% to both  $\tilde{\mu}_L \nu_\mu$  and  $\tilde{e}_L \nu_e$ . Despite the low cross section there is a high number of trilepton events.
- Point 8:  $\sigma_{p8} = 4.6$  pb,  $m_{\tilde{\chi}_1^0} = 56$  GeV,  $m_{\tilde{\chi}_2^0, \tilde{\chi}_1^+} = 200$  GeV,  $m_{\tilde{e}_L} = 107$  GeV and  $m_{\tilde{\mu}_L} = 102$  GeV. The left-handed sleptons are lightest while the right-handed ones have a mass of approximately 465 GeV.  $\tilde{\chi}_2^0$  has the highest branching fraction to  $\tilde{\nu}_{e,\mu} \nu_{e,\mu}$ . However, approximately at the same level,  $\tilde{\chi}_2^0$  decays also to  $\tilde{\mu}_L \mu$  and  $\tilde{e}_L e$  (24%, 22% respectively).  $\tilde{\chi}_1^+$  decays mainly via a sneutrino plus a lepton ( $\sim 31\%$  per flavor,  $e, \mu$ ). The branching fraction to light left-handed slepton plus lepton is approximately 17% per flavor ( $e, \mu$ ).

The listing above focuses on  $\tilde{\chi}_2^0$  and  $\tilde{\chi}_1^+$  as production of such pairs have the largest cross section. Detailed information about the points can be found in [177]

|                       | ev. after Cut 6 | $S_0$ | $Z_{LLR}$ | $S_N$ | $Z_N$ |
|-----------------------|-----------------|-------|-----------|-------|-------|
| directgaugino_sle.p01 | 9.0             | 7.5   | 5.7       | 4.8   | 3.2   |
| directgaugino_sle.p02 | 0.7             | 0.5   | 1.0       | 0.4   | 0.3   |
| directgaugino_sle.p03 | 3.7             | 3.1   | 3.1       | 2.0   | 1.5   |
| directgaugino_sle.p04 | 1.6             | 1.4   | 1.7       | 0.9   | 0.7   |
| directgaugino_sle.p05 | 5.3             | 4.4   | 4.0       | 2.8   | 2.0   |
| directgaugino_sle.p06 | 4.1             | 3.4   | 3.3       | 2.2   | 1.6   |
| directgaugino_sle.p07 | 6.9             | 5.8   | 4.8       | 3.7   | 2.6   |
| directgaugino_sle.p08 | 9.5             | 8.0   | 5.9       | 5.1   | 3.3   |
| directgaugino_sle.p09 | 2.8             | 2.3   | 2.5       | 1.5   | 1.2   |
| directgaugino_sle.p10 | 3.7             | 3.1   | 3.1       | 2.0   | 1.5   |
| directgaugino_sle.p11 | 1.4             | 1.1   | 1.6       | 0.7   | 0.6   |
| directgaugino_sle.p12 | 2.1             | 1.7   | 2.0       | 1.1   | 0.9   |
| Sum background        | 1.4             |       |           |       |       |

Table 10.3: Significances for the jet exclusive search at 10 TeV, direct-gaugino points with decay via sleptons ( $\tilde{e}, \tilde{\mu}$ ). All numbers are normalized to 200 pb<sup>-1</sup>. 100% background uncertainty.

|                  | ev. after Cut 6 | $S_0$ | $Z_{LLR}$ | $S_N$ | $Z_N$ |
|------------------|-----------------|-------|-----------|-------|-------|
| directgaugino.p1 | 4.6             | 3.8   | 3.6       | 2.4   | 1.8   |
| directgaugino.p2 | 3.8             | 3.2   | 3.1       | 2.0   | 1.5   |
| directgaugino.p3 | 0.1             | 0.1   | 0.4       | 0.1   | 0.0   |
| directgaugino.p4 | 0.1             | 0.1   | 0.3       | 0.0   | 0.0   |
| Sum background   | 1.4             |       |           |       |       |

Table 10.4: Significances for the jet exclusive search at 10 TeV, direct-gaugino points at. All numbers are normalized to 200 pb<sup>-1</sup>, 100% background uncertainty. Track-based isolation  $p_{T\Sigma\text{track}}^{\Delta R=0.3}/p_T < 0.11$ .

Scenarios like point 1, 8 and possibly also 7 have around  $5\sigma$  discovery potential even with a relatively small data sample.

The direct-gaugino-stau points where gauginos may decay via a stau have very low significances and will not be discussed in detail. With 200 pb<sup>-1</sup> even the most optimistic significance measure does not exceed one. Such scenarios are interesting, but will require larger data samples. The optimistic  $S_0$  shows that one needs at least  $\sim 15$  fb<sup>-1</sup> to reach  $5\sigma$ .

### 10.1.2 Alternative track based isolation

Since the additional isolation requirement in Cut 3 reduces strongly the signal, it is interesting to try as an alternative the relative track based isolation  $p_{T\Sigma\text{track}}^{\Delta R=0.3}/p_T$ . The 14 TeV analysis showed that this cut provides a strong rejection of *extra* leptons while still keeping a good fraction of the *prompt* ones. As this is only meant as a test, there has not been any new optimization study involving the new SM backgrounds.

Cut 3 in Section 10.1.1 is now replaced by the requirement  $p_{T\Sigma\text{track}}^{\Delta R=0.3}/p_T < 0.11$ . All

|                       | ev. after Cut 6 | $S_0$ | $Z_{LLR}$ | $S_N$ | $Z_N$ |
|-----------------------|-----------------|-------|-----------|-------|-------|
| directgaugino_sle.p01 | 9.7             | 8.1   | 6.0       | 5.2   | 3.3   |
| directgaugino_sle.p02 | 0.7             | 0.5   | 1.0       | 0.4   | 0.3   |
| directgaugino_sle.p03 | 4.2             | 3.5   | 3.3       | 2.2   | 1.7   |
| directgaugino_sle.p04 | 1.6             | 1.4   | 1.7       | 0.9   | 0.7   |
| directgaugino_sle.p05 | 5.2             | 4.4   | 3.9       | 2.8   | 2.0   |
| directgaugino_sle.p06 | 4.5             | 3.8   | 3.5       | 2.4   | 1.8   |
| directgaugino_sle.p07 | 8.6             | 7.2   | 5.5       | 4.6   | 3.0   |
| directgaugino_sle.p08 | 11.2            | 9.3   | 6.6       | 6.0   | 3.7   |
| directgaugino_sle.p09 | 3.7             | 3.0   | 3.0       | 2.0   | 1.5   |
| directgaugino_sle.p10 | 4.8             | 4.0   | 3.7       | 2.6   | 1.9   |
| directgaugino_sle.p11 | 1.9             | 1.6   | 1.9       | 1.0   | 0.8   |
| directgaugino_sle.p12 | 2.7             | 2.2   | 2.4       | 1.4   | 1.1   |
| Sum background        | 1.4             |       |           |       |       |

Table 10.5: Significances for the jet exclusive search at 10 TeV, direct-gaugino points with decay via sleptons ( $\tilde{e}, \tilde{\mu}$ ). All numbers are normalized to 200 pb<sup>-1</sup>, 100% background uncertainty. Track-based isolation  $p_{T\Sigma\text{track}}^{\Delta R=0.3}/p_T < 0.11$ .

other event selection cuts remain unchanged. Table 10.4 and 10.5 list the number of events after all cuts and the significances for the direct-gaugino and direct-gaugino-slepton points respectively.

For almost all points there is a small improvement. With the exception of the conservative significance measure  $Z_N$ , all the other significance measures show for points 1, 7 and 8 values of 5 or more. This shows once gain that an optimized isolation cut can improve the jet exclusive trilepton analysis and make it interesting even with relatively small data samples.

## 10.2 Gauginoshake

The patterns of the mSUGRA decay chains are defined by the typical gaugino composition where the LSP is pure bino,  $\tilde{\chi}_1^\pm$  and  $\tilde{\chi}_2^0$  dominantly wino and the remaining heavier gauginos mostly higgsino like. While in the gauginoshake points the hierarchy of  $M_1$ ,  $M_2$  and  $\mu$  is changed such that:

- point 1 the LSP is dominantly higgsino,
- point 2 the LSP is dominantly wino
- point 3 the LSP is a full mix.

Also the composition of the heavier gauginos may be very different compared to mSUGRA. What also makes these points interesting is that they represent low mass SUSY models favoured by the current experimental conditions. The only difference from conventional mSUGRA models lies in the composition of the gauginos.

As listed in Table 10.6, the three points in this category have low squark and gluino masses and therefore relatively high cross sections (16-26 pb) dominated by  $\tilde{q}\tilde{q}$  and  $\tilde{q}\tilde{q}$

| Sample         | $\sigma_{LO}[\text{pb}]$ | $m_{\tilde{\chi}_1^0} [\text{GeV}]$ | $m_{\tilde{\chi}_2^0}$ | $m_{\tilde{\chi}_1^\pm}$ | $m_{\tilde{\tau}_1}$ | $m_{\tilde{\mu}/\tilde{e}}$ | $m_{\tilde{u}_L}$ | $m_{\tilde{g}}$ |
|----------------|--------------------------|-------------------------------------|------------------------|--------------------------|----------------------|-----------------------------|-------------------|-----------------|
| gauginoshake 1 | 25.8                     | 85                                  | 107                    | 92                       | 250                  | 157                         | 497               | 600             |
| gauginoshake 2 | 16.1                     | 184                                 | 363                    | 189                      | 239                  | 254                         | 497               | 600             |
| gauginoshake 3 | 18.6                     | 134                                 | 200                    | 144                      | 195                  | 156                         | 497               | 600             |

Table 10.6: Overview of the MSSM24 gauginoshake points. All masses are in GeV. [177]

production. In point 1 where the  $\tilde{\chi}_1^+$  and  $\tilde{\chi}_2^0$  are relatively light, pair production of these account for about 30% of the cross section.

In all three points the gluino has a mass of 600 GeV, the left-handed squarks approximately a mass of 500 GeV, while the right-handed squarks approximately 550 GeV. The mass splitting in the third family is noticeable only in point 3 where  $\tilde{b}_1$  and  $\tilde{t}_1$  are lightest (403 GeV, 410 GeV respectively). In point 1 and 2 the gluino has the highest decay rate to the left-handed squarks  $\tilde{g} \rightarrow \tilde{q}_L q$  of around 15% per flavor, except for  $\tilde{t}_1$  which is too heavy. In point 3 where  $\tilde{b}_1$  and  $\tilde{t}_1/\tilde{b}_2$  are lighter they account for 26% and 17%/16%, respectively.

The squarks decay in many cases to the heavier gauginos due to their less conventional composition as compared to mSUGRA. In such cases they initiate a longer decay chain which can involve lighter gauginos or sleptons which again contribute to high lepton multiplicity in the final state. With the accompanying high jet activity and significant  $\cancel{E}_T$ , these points are very suitable for a jet inclusive analysis type. Figure 10.5 shows the number of leptons (a). It is interesting to note that point 1 and 3 have a high number of events with 4 leptons. Figure 10.5 (b) shows the lepton  $p_T$  after requiring  $N_l \geq 3$  and  $M_{l+l-} > 20$  GeV. Especially point 1 and 3 have a significantly harder lepton  $p_T$  spectrum than the SM background. In addition, in plot (c) one can see the relatively high number of jets having  $p_T > 10$  GeV.

### 10.2.1 Jet inclusive analysis

The gauginoshake points have such outstanding signatures that they would be easy to see even in an analysis which is not optimized for such scenarios.

The object definitions are the same as described in Section 9.3 and the event selection is based on the cuts from Section 9.5:

1.  $p_T^{jet\ 1} > 80$  GeV,  $p_T^{jet\ 2} > 50$  GeV
2.  $\cancel{E}_T > 110$  GeV

Figure 10.6 shows the event selection variables used in the jet inclusive searches developed for the SUx points in Section 9.5:  $p_T$  of the leading (a) and second leading jet (b), and  $\cancel{E}_T$  (c).

This event selection is optimized for an SU4 like scenario with, in an mSUGRA context at least, soft jets and moderate  $\cancel{E}_T$ . An event selection more suitable for scenarios like the gauginoshake points should impose harder jet  $p_T$  requirements, possibly more than two jets, and harder  $\cancel{E}_T$  cut. Nevertheless, from Table 10.7, which lists the significances for the various points, one can see that all points have promising discovery potential. With the exception of the conservative  $Z_N$  significance measure for point 2, all points yield a  $5\sigma$  discovery even with  $200 \text{ pb}^{-1}$ .

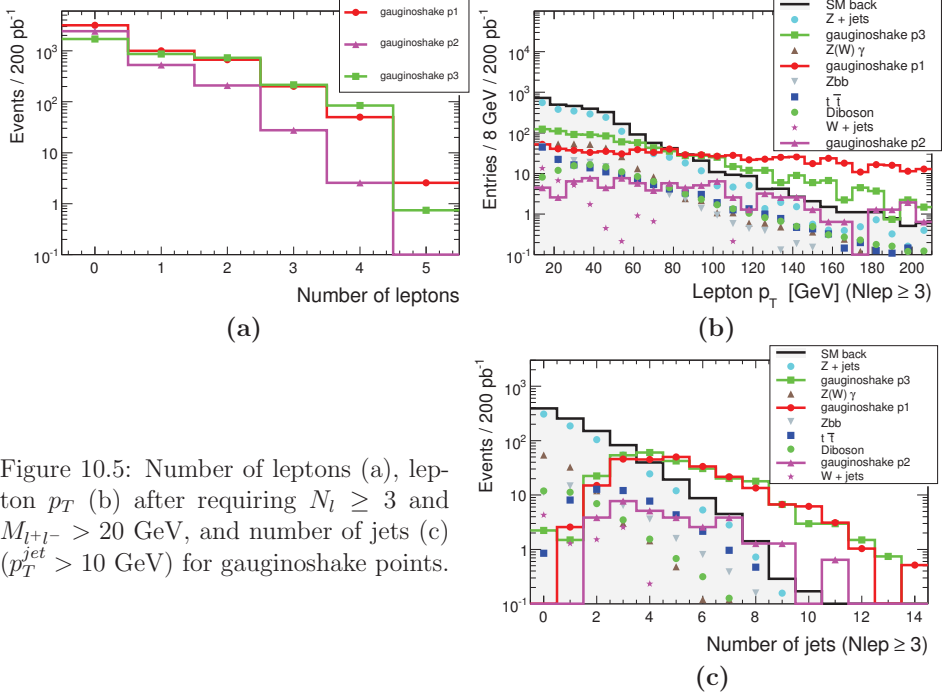


Figure 10.5: Number of leptons (a), lepton  $p_T$  (b) after requiring  $N_l \geq 3$  and  $M_{l+l-} > 20$  GeV, and number of jets (c) ( $p_T^{jet} > 10$  GeV) for gauginoshake points.

|                 | ev. after Cut 2 | $S_0$ | $Z_{LLR}$ | $S_N$ | $Z_N$ |
|-----------------|-----------------|-------|-----------|-------|-------|
| gauginoshake_p1 | 139.7           | 87.1  | 31.7      | 46.1  | 18.6  |
| gauginoshake_p2 | 18.0            | 11.2  | 8.2       | 5.9   | 4.1   |
| gauginoshake_p3 | 176.2           | 109.8 | 36.6      | 58.1  | 22.1  |
| Sum background  | 2.6             |       |           |       |       |

Table 10.7: Significances jet inclusive search at 10 TeV, gauginoshake points at 10 TeV. All numbers are normalized to 200 pb<sup>-1</sup>. 100% background uncertainty.

### 10.2.2 4-lepton signal

The general trilepton analysis requires three or more leptons. However, scenarios like gauginoshake have shown in Figure 10.5 (a) a considerable fraction of events with four leptons. This combined with the relatively high cross section makes it interesting also to explore the potential of four lepton analysis.

Requiring four leptons leads to a very strong background suppression. The additional spectacular jet and  $\cancel{E}_T$  characteristics make it possible to obtain a background free signal even with relatively moderate cuts.

Figure 10.7 show 2D significance scans using  $Z_N$  with 100% background uncertainty as significance measure. The scan is showed for the two points with highest fraction of four lepton events, gauginoshake point 1 and 3. Panels (a) and (c) show  $Z_N$  as function of the cut on  $p_T$  of the two leading jets. While panels (b) and (d) show it as function of the cuts on  $p_T$  of the leading jet and  $\cancel{E}_T$ . As both these points have an extraordinary high  $\cancel{E}_T$  spectrum, combined with a cut on the leading jet leads to the highest significances. The



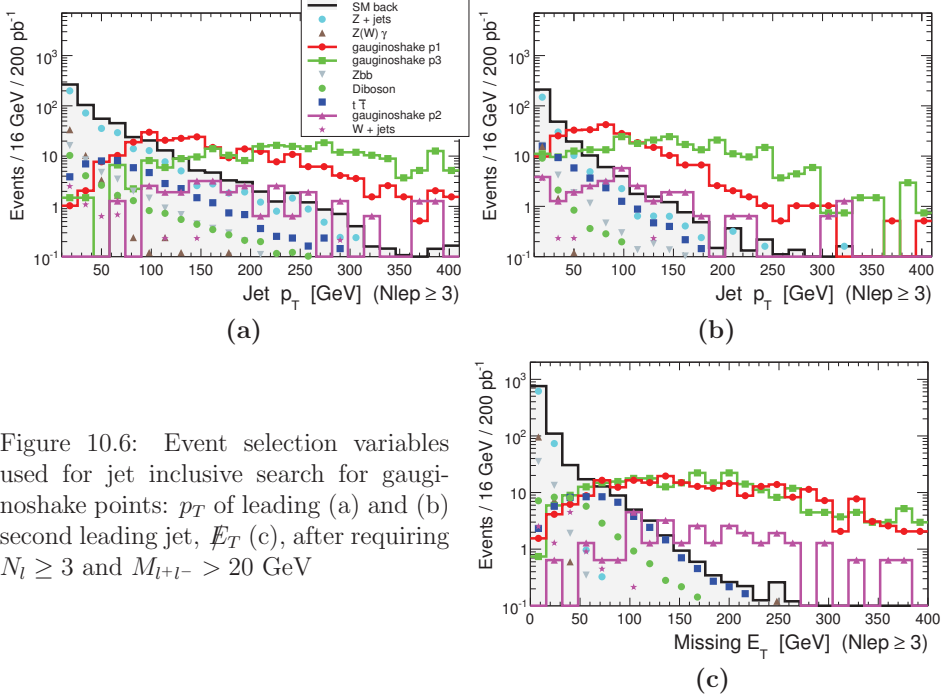


Figure 10.6: Event selection variables used for jet inclusive search for gauginoshake points:  $p_T$  of leading (a) and (b) second leading jet,  $\cancel{E}_T$  (c), after requiring  $N_l \geq 3$  and  $M_{l+l-} > 20$  GeV

uniform red/dark area in the upper right part of the plots correspond to a “background free” region which has been defined as  $B < 0.00001$  and  $S > 2$ . The combination of a cut on the leading jet  $p_T$  and  $\cancel{E}_T$  leads to higher maximum significances than a combined cut in the two leading jets.

The three dimensional scan which was earlier used to find optimized cuts for the jet inclusive SU4 search has been applied on the 4-lepton gauginoshake signal and SM background. The scan varies simultaneously the minimum cut on  $p_T$  of the two leading jets and  $\cancel{E}_T$ . In this case it explores the  $p_T^{jet\ 1} - p_T^{jet\ 2} - \cancel{E}_T$  space with no background left. The optimized quantity is the number of signal events. Table 10.8 shows the cuts which lead to highest number of events.

The jet inclusive search optimized for the trilepton SU4 signal was also applied requiring four leptons (same object definitions as in Section 9.3). Requiring a high number of leptons reduces strongly the background and even the relatively moderate cuts lead to high significances for point 2 and 3. Table 10.9 summarizes the results. With this set of cuts there is almost no SM background left, while the number of signal events is significant, especially for points 1 and 3.

This preliminary study shows that a four lepton requirement with moderate jet  $p_T$  and  $\cancel{E}_T$  cuts is for some SUSY scenarios a very interesting discovery channel already during early data taking. Although the set of cuts involves both jets and  $\cancel{E}_T$  one would rely mostly on leptons which are expected to be better understood at an early point.

|                 | $p_T^{jet\ 1}$ [GeV] | $p_T^{jet\ 2}$ [GeV] | $\cancel{E}_T$ [GeV] | Num. events. |
|-----------------|----------------------|----------------------|----------------------|--------------|
| gauginoshake.p1 | 140.0                | 10.0-30.0            | 100.0                | 11.3         |
| gauginoshake.p2 | 150.0                | 70.0-100.0           | 30.0-40.0            | 1.9          |
| gauginoshake.p3 | 160.0                | 30.0                 | 50.0                 | 43.1         |

Table 10.8: Cut scan optimizing the number of 4-lepton gauginoshake signal events in the  $p_T^{jet\ 1} - p_T^{jet\ 2} - \cancel{E}_T$  space with no SM background left. All numbers are normalized to 200 pb<sup>-1</sup>.

|                 | ev. after Cut 2 | $S_0$  | $Z_{LLR}$ | $S_N$  | $Z_N$ |
|-----------------|-----------------|--------|-----------|--------|-------|
| gauginoshake.p1 | 25.3            | 1032.7 | 22.7      | 1032.4 | 20.5  |
| gauginoshake.p2 | 0.6             | 26.2   | 2.9       | 26.2   | 2.4   |
| gauginoshake.p3 | 46.1            | 1884.5 | 31.5      | 1884.0 | 28.3  |
| Sum background  | 0.0006          |        |           |        |       |

Table 10.9: Significances for jet inclusive search at 10 TeV requiring four leptons, gauginoshake points. All numbers are normalized to 200 pb<sup>-1</sup>.

## 10.3 Conclusions

The majority of SUSY studies in ATLAS have so far been based on mSUGRA models which, with few exceptions, share common phenomenological characteristics. In this section we investigated how the analysis strategies developed for mSUGRA are sensitive to different scenarios.

Both the jet inclusive and exclusive analyses have been applied to a wide range of new SUSY scenarios defined in the MSSM24 PhenoGrid framework. The earlier developed jet exclusive strategy is suitable for a larger range of models where the gaugino pair production is the only possible or dominant production process. Some of the studied scenarios show good discovery potential even with a relatively small data sample like 200 pb<sup>-1</sup>.

The jet inclusive analysis is also shown to be powerful in the context of new multi lepton and multi jet scenarios. Especially a requirement of four leptons may in such scenarios lead to very high discovery potential.

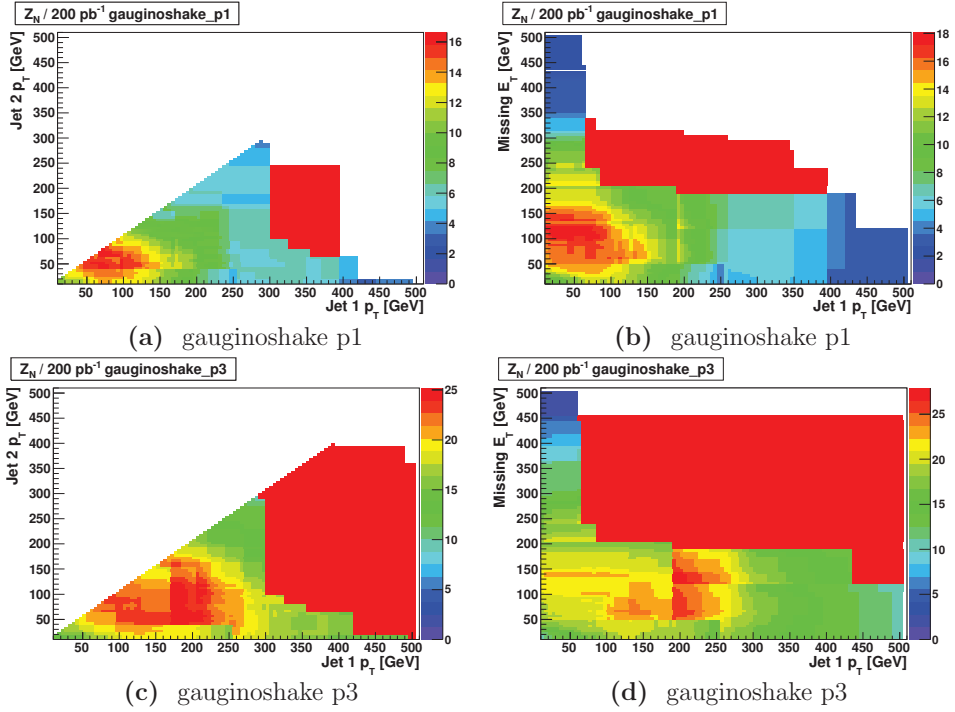


Figure 10.7: Gauginoshake point 1 and 3 significance scan requiring 4 leptons.  $Z_N$  is shown as function of cuts on the  $p_T$  of the two leading jets (a),(c) and as function of the cut on  $p_T$  of the leading jet and  $\cancel{E}_T$ (b),(d). The uniform red area in the upper right part corresponds to “background free” signal defined as  $B < 0.00001$  and  $S > 2$ .



# Chapter 11

## First collision data

Towards the end of November 2009 LHC started circulating beams after recovering from the helium blowout which occurred 19th September 2008. After only a couple of days the beams were under good enough control to be collided. The first collision candidate in ATLAS was observed on the 23rd of November 2009. During the remaining weeks of 2009, LHC continued to deliver collisions at 900 GeV center of mass energy which corresponds to the injection energy from the SPS. Towards the end of the data taking in 2009, the energy was increased to 2.36 TeV which is a new world record.

This section presents a first look at collision data. The goal is not to study in detail the performance of the ATLAS detector. It is rather to compare the physics objects which were used in the MC-based analyses with data in order to see how well the simulation reproduces it. The isolation variables which are important for the trilepton analysis will be presented more extensively.

### 11.1 Event selection

The data included in this analysis are defined as “good runs” at the center of mass energy of 900 GeV. One of the selection criteria is that the solenoid magnet is on. In runs recommended for muon studies also the toroid field is required to be on. Analyzed events are from luminosity blocks <sup>1</sup> for which the condition data base indicates that important parts of the inner detector and calorimeters were operational. In addition, the beam condition is required to be stable. The data collection used in this analysis has been through one reprocessing after the first reconstruction <sup>2</sup>. The reprocessing implemented a number of software improvements notably in tracking, muon reconstruction. It used also an updated condition database. A set of MC simulations which correspond to the reprocessed data was also produced [178].

The following list of  $\sqrt{s} = 900$  GeV runs was analysed:

- For electron study: 141749, 141811, 142149, 142154, 142165, 142166, 142171, 142174, 142189, 142191, 142193, 142195, 142383
- For muon study, runs with toroid on: 142174, 142189, 142191, 142193, 142195, 142383

---

<sup>1</sup>The list of “good runs” and accepted luminosity blocks is the one provided by the electron-photon performance group and the muon group.

<sup>2</sup>tag: r988.p62, DESD\_COLL CAND files which are derived ESDs collision candidates

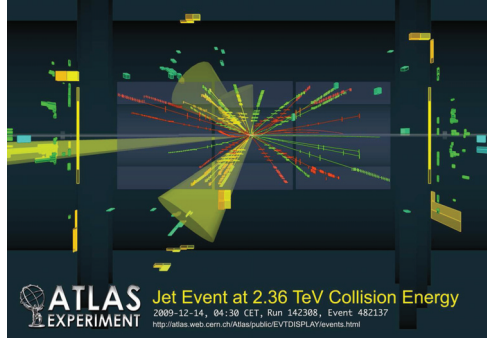


Figure 11.1: ATLAS event display of a jet event at 2.36 TeV.

|   | Event selection      | Number of events | Efficiency [%] |
|---|----------------------|------------------|----------------|
| 1 | 1-lep. cand. ev.     | 3816             |                |
| 2 | MBTS timing side A+C | 3776             | 99.0           |
| 3 | MBTS_1_1 Trigger     | 3801             | 99.6           |
| 4 | Primary vertex       | 3763             | 98.6           |
| 5 | All requirements     | 3740             | 98.0           |

Table 11.1: Breakdown of minimum bias event selection requirements. Row 2-4 give the individual requirement efficiencies, row 5 the combined efficiency.

Real data was compared to corresponding  $\sqrt{s} = 900$  GeV MC <sup>3</sup>

The 2.36 TeV samples are very small and as no leptons passed the selection cuts, they are not included in the analysis. Figure 11.1 shows an ATLAS event display of a di-jet event at 2.36 TeV.

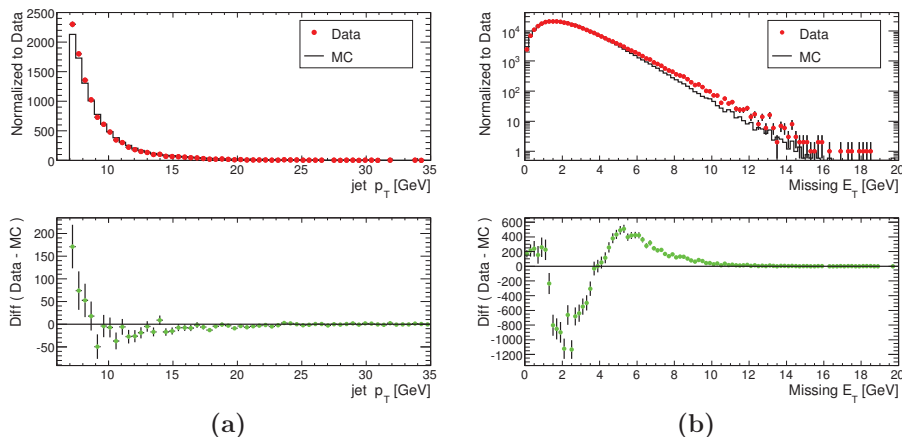
As the focus is again on leptons, the analysis was performed on D2PDs which only contain events with one or more lepton candidates.

In order to pick out real minimum bias collisions an event has to pass three requirements. First it must pass the minimum bias collision requirements provided by the Minimum Bias Trigger Scintillator (MBTS) [73] which is mounted on the LAr endcap cryostat perpendicular to the beam line. It is required that the difference between the average hit time at side A and C<sup>4</sup> is less than 10 ns and that the event is accepted by the Level 1 MBTS\_1\_1 (L1\_MBTS\_1\_1) trigger. An event is accepted only if it is triggered both by side A and C. In addition it is required that the event has a reconstructed primary vertex.

Before any event selection cuts, a sample of 3816 events is obtained in the selected runs and luminosity blocks with at least one lepton candidate. After the three event selection requirements there are 3740 events left which gives an efficiency of 98.0%. Table 11.1 shows a breakdown of the efficiencies for the three event selection criteria described above. Row 2-4 give the efficiencies of the individual requirements, while row 5 shows the combined efficiency of the three minimum bias requirements.

<sup>3</sup>mc09\_900GeV.105001.pythia\_minbias.recon.AOD.e500\_s655\_s657\_d257\_r1023, reprocessing with updated magnetic field

<sup>4</sup>ATLAS is divided in three parts: the barrel B and the two endcap sides A and C.

Figure 11.2: Comparison data-MC jet  $p_T$  (a) and  $E_T$  (b).

## 11.2 Jets and missing $E_T$

Energetic jets and considerable  $E_T$  are expected to be some of the most striking characteristics of many SUSY scenarios. It will therefore be important to obtain good understanding of the calorimeters, calibration and jet energy scale, as well as of the jet reconstruction. These issues are known to be difficult and the uncertainties related to jets and  $E_T$  are expected to be large during the early data taking.

Studies done by the jet/ $E_T$  performance group show that requirements of a minimum fraction of electromagnetic energy deposit in jets and in the event as a whole rejects events with large contributions of fake  $E_T$ . These so-called cleaning cuts remove also fake  $E_T$  arising from cosmic muons [179].

The jet candidates in this analysis are taken from the Anti- $K_T$  topological jet collection<sup>5</sup> and were required to have an electromagnetic energy fraction ( $J_{EMfrac}$ ) larger than 0.1. Their  $p_T$  distribution is shown in Figure 11.2 (a) which contains all events, not only 1-lepton events. From the lower plot of Figure (a) which shows the difference between data and MC as a function of jet  $p_T$ , one can see that there is a good agreement. However, there are clearly more jets in the low  $p_T$  region up to about 10 GeV in data than in MC.

Figure 11.2 (b) shows the distributions of the  $E_T$  variable called MET\_Reffinal (see Section 7.1). It is the same  $E_T$  summation as used in the analyses presented earlier. An event electromagnetic energy fraction is defined as:

$$EV_{EMfrac} = \frac{\sum J_{EMfrac} \times E_T^{Jet}}{\sum E_T^{Jet}} \quad (11.1)$$

where  $E_T^{Jet}$  is the transverse jet energy and the sum runs over all jets in the event. The event electromagnetic energy fraction,  $EV_{EMfrac}$ , was required to be larger than 0.1. The  $E_T$  distribution in data and MC are close, however even after the cleaning cut, there are more events in the tail in data and from the lower plot one can see that the mean value is higher than in MC.

<sup>5</sup>The topological cluster cells used in reconstruction include noise suppression and calibration which is an advantage during this early data taking.

## 11.3 Tracks

In ATLAS a number of track objects are related to the different sub-detector systems and reconstruction algorithms. The objects used here belong to the common physics analysis TrackParticleCandidate class which represents inner detector tracks. The following track selection requirements are imposed:

- Transverse impact parameter  $|d_0| < 1.0$  mm, longitudinal impact parameter  $|z_0 \sin \theta| < 1.5$  mm. The impact parameters are calculated with respect to the reconstructed vertex. These cuts ensure that tracks originate from the primary vertex and reject background tracks, for example from cosmic rays and beam gas interactions. Figure 11.3 gives a graphical definition of the impact parameters.
- Track quality cuts: Number of Pixel detector hits  $N_{pix} \geq 3$ , number of SCT hits  $N_{SCT} \geq 6$ . The requirement ensures that the track reconstruction is based on a certain number of measurements in order to keep only good quality tracks.
- Phase-space:  $p_T > 500$  MeV and  $|\eta| < 2.6$ .

Figure 11.4 shows a comparison of the distribution of number of tracks (a), track  $p_T$  (b),  $\eta$  (c) and  $\phi$  (d). All distributions include tracks that have passed the track selection and include all events, not only 1-lepton events.

Comparing the number of tracks (a) MC does not reproduce the low track multiplicity and the whole distribution is shifted towards higher values. This can indicate that there are more tracks in MC that pass the track selection criteria. The  $p_T$  (b) and  $\phi$  (d) distributions agree well, while data  $\eta$  distribution show a significant asymmetry about zero which is not matched by MC. This effect can be related to the vertex displacement which is not fully reproduced in the simulation.

Figure 11.5 shows the number of tracks (a) and track  $p_T$  (b) for events with one or more lepton candidates which passed the selection criteria above. The lower panels show the difference of data and MC. There is a relative good agreement, except for the first bin in figure (b) representing tracks with a  $p_T$  close to the cut value of 0.5 GeV. The jump at 2 GeV reflects the fact that one lepton is required and some of the reconstruction algorithms provide candidates with a minimum  $p_T$  of 2 GeV.

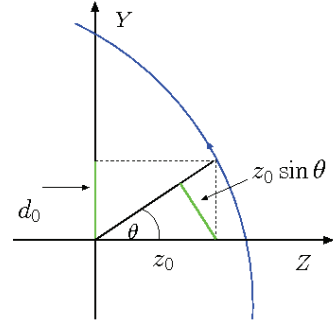


Figure 11.3: Impact parameters  $d_0$ ,  $z_0$  and  $z_0 \sin \theta$ .

## 11.4 Muons

Muon candidates are as in the earlier analyses taken from the Staco collection [163]. The selection criteria are:

- Transverse impact parameter  $|d_0| < 1.0$  mm, longitudinal impact parameter  $|z_0 \sin \theta| < 1.5$  mm. The impact parameters are calculated with respect to the reconstructed vertex.



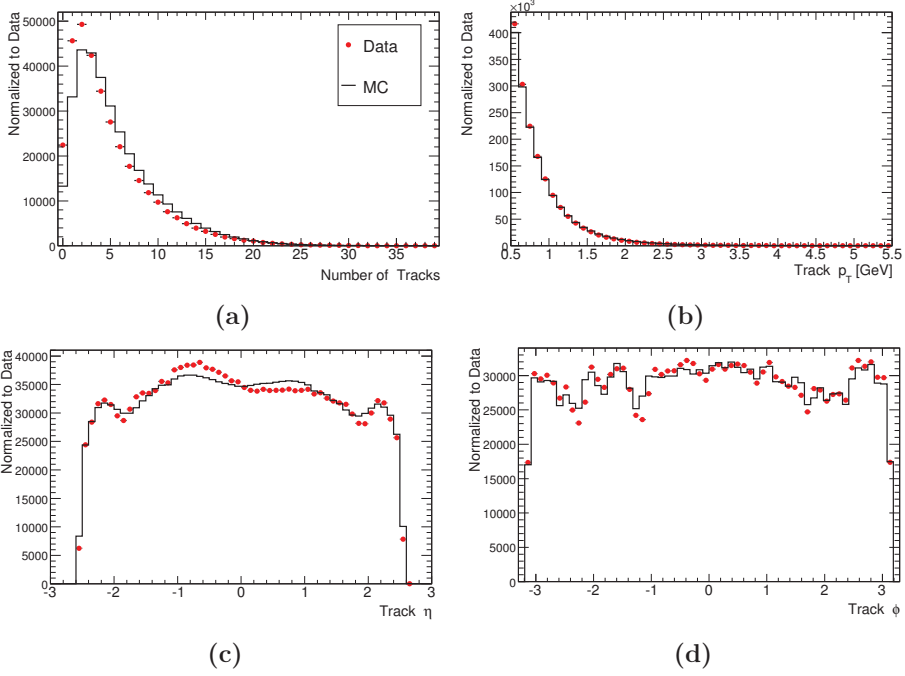


Figure 11.4: Comparison data - MC: (a) number of tracks per event, (b)  $p_T$ , (c)  $\eta$ , and (d)  $\phi$ .

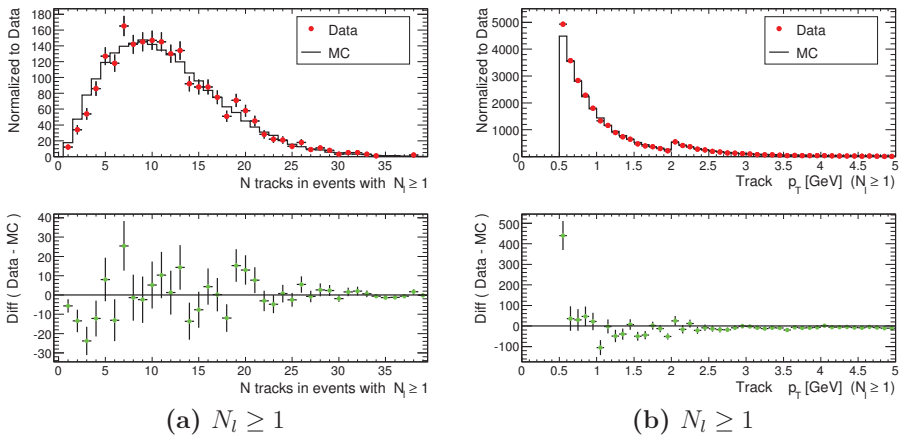


Figure 11.5: Comparison data-MC for events with  $N_l \geq 1$ : (a) number of tracks and (b) track  $p_T$ .

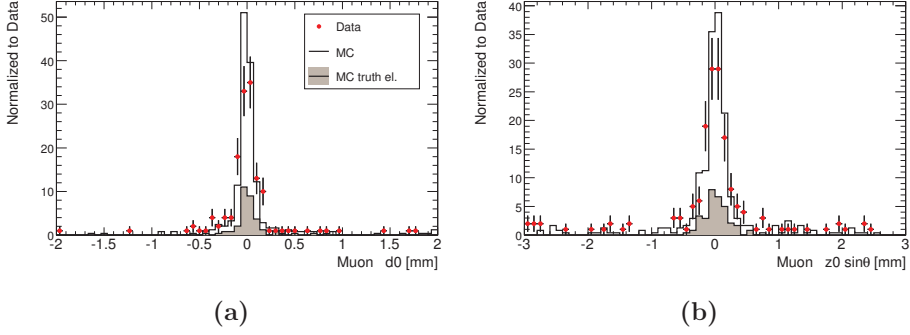


Figure 11.6: Comparison data-MC: (a) the transverse impact parameter  $d_0$  and (b) the longitudinal impact parameter  $z_0 \sin \theta$  measured with respect to the reconstructed primary vertex. No track quality cuts are applied, combined and low  $p_T$  muons.

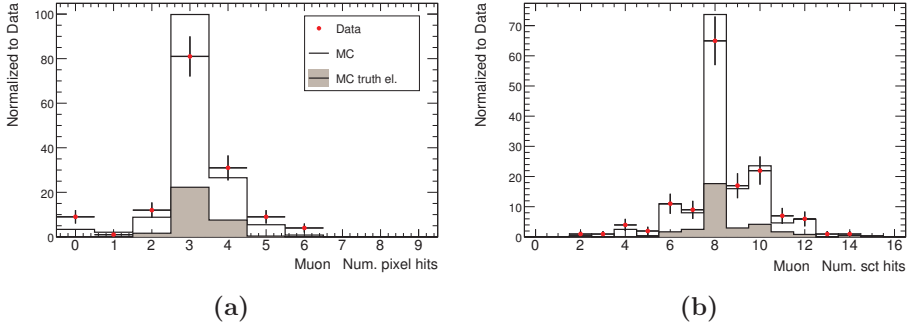


Figure 11.7: Comparison data-MC: (a) number of Pixel hits and (b) SCT hits for muons. No impact parameter cuts are applied, combined and low  $p_T$  muons.

- Track quality cuts: Number of pixel detector hits  $N_{pix} \geq 3$ , number of SCT hits  $N_{SCT} \geq 6$ .
- $p_T > 500$  MeV and  $|\eta| < 2.5$ .
- A combined muon, which is reconstructed from tracks both in the inner detector and Muon Spectrometer, is required to have a matching  $\chi^2 < 100$ . In addition, the study includes also muons tagged as low- $p_T$  muons. These are reconstructed using specialized algorithms which start from an inner detector track and extrapolate it to the muon spectrometer. Due to low momentum, they have only a track segment in the first layer of the spectrometer [161].

The impact parameter cuts are important in order to remove muons from cosmic radiation background. Figure 11.6 (a) shows the transverse impact parameter  $d_0$  while (b) shows the longitudinal impact parameter  $z_0 \sin \theta$  before any track quality cuts are applied. There is some agreement between simulation and data, although both the  $d_0$  and  $z_0 \sin \theta$  distribution in MC are clearly narrower than in data. The shaded histogram is the distribution for MC muons that are matched to a MC truth muon within  $\Delta R = 0.03$ . Clearly, although one speaks of muons the majority of candidates are fakes which can be associated to pions or kaons.

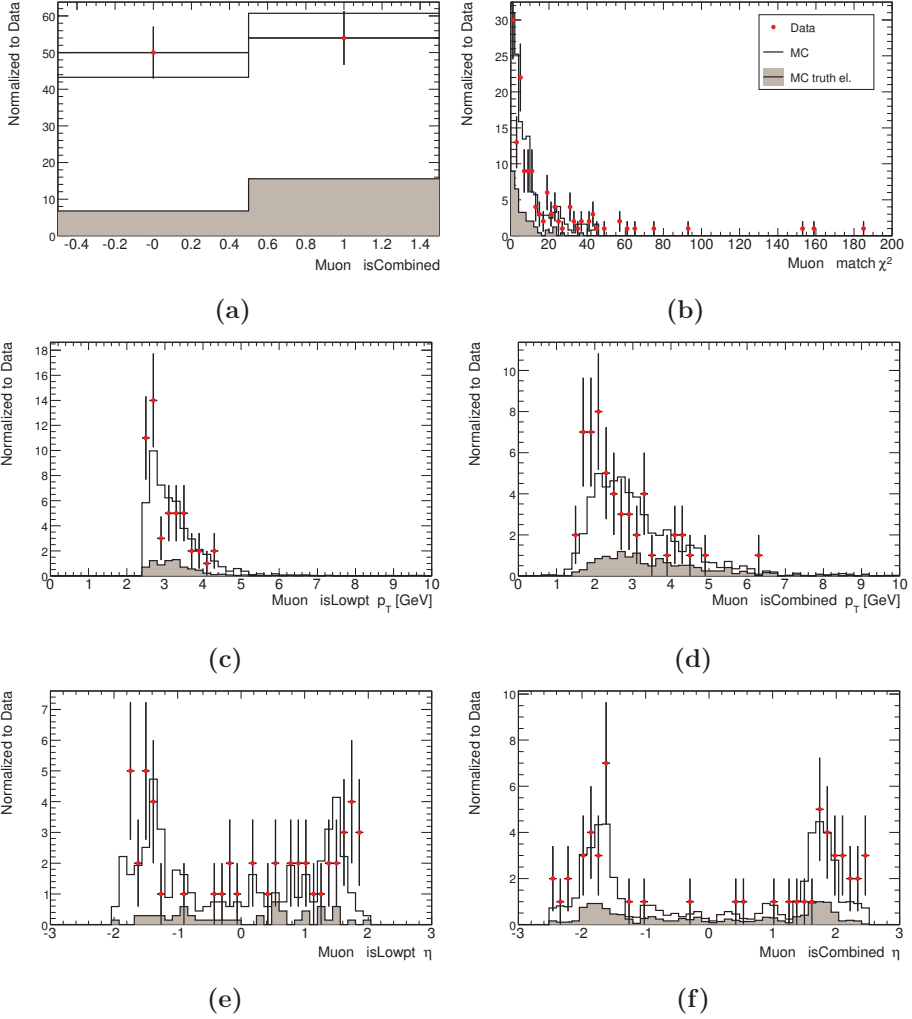


Figure 11.8: Comparison data - MC: (a) combined and low  $p_T$  muon tags, (b) track match  $\chi^2$  for combined muons, (c)-(d)  $p_T$  for low- $p_T$  and combined muons and (e)-(f)  $\eta$  low- $p_T$  and combined muons

Requiring a certain number of hits in the inner detector improves greatly the track quality. From Figure 11.7, which shows the number of Pixel (a) and SCT (b) hits, we see that the applied cuts ( $N_{pix} \geq 3$ ,  $N_{SCT} \geq 6$ ) keep the majority of the muon candidates and there is a reasonable agreement between data and MC.

Muons are tagged according to the reconstruction algorithm, either as combined or low- $p_T$  muons. Figure 11.8 (a) shows that in MC a higher fraction of muons is tagged as combined and correspondingly lower fraction is tagged as low- $p_T$  muons. However, the discrepancy is within the statistical error. A combined muon is a result of a matching of an inner detector track and a Muon Spectrometer track. The  $\chi^2$  related to the match is shown in Figure 11.8 (b). In MC the distribution is shifted towards lower values and does not reproduce well the long tail. Following the selection criteria from the 14 and 10 TeV analysis, the match  $\chi^2$  is required to be less than 100.

The  $p_T$  distributions for the low  $p_T$  and combined muons are shown in Figure 11.8 (c)-(d), respectively. In the MC distributions, both for combined and low  $p_T$  muons, there are clearly less entries in the first bins representing the lowest  $p_T$ . As the statistics is low it is difficult to compare the  $\eta$  distributions. However, for both classes, at high  $|\eta|$  values the peaks in MC distributions seem to be shifted towards lower  $|\eta|$  values.

After applying all muon selection cuts, there are 91 combined and 69 low  $p_T$  muon candidates in the analysed runs.

#### 11.4.1 Muon isolation

In the current low energy runs there are no real isolated muons of the type that were focused on in the SUSY analyses (Chapter 8,9 and 10). Most of the candidates are pions or kaons which are decay products of  $B$  and  $D$  mesons. However, it is still relevant to study the various isolation variables in order to see how well the MC fits with data. The distributions in this section include both combined and low- $p_T$  muons in order to obtain higher statistics.

The most commonly used calorimeter-based isolation variable  $E_{cal}^{\Delta R=0.X}$  (see Section 7.2.2) is shown in Figure 11.9 for all muon candidates. Despite the low statistics, the simulated distribution follows well the data.

The true MC muons represented by the shaded histogram all originate from semileptonic heavy flavor decay. The distributions seem slightly narrower, however, in high energy collisions one expects these kind of muons to be less isolated than those from  $W$ ,  $Z$ ,  $\tau$  or SUSY decays.

Figure 11.10 shows the track-based isolation variable  $p_{T\Sigma track}^{\Delta R=0.X}$  (see Section 7.2.2) which is the a sum of track  $p_T$  in a cone around the muon track (default variables associated with the muon object). Tracks that are allowed in the sum are required to have a minimum  $p_T$  of 1.0 GeV and given the soft spectrum of tracks in the 900 GeV collisions, only a few muon candidates have  $p_{T\Sigma track}^{\Delta R=0.X} > 1$ . GeV. For low energy runs one might consider a lower minimum cut on tracks in the sum in order to be more sensitive to the activity around the muons. One can see that the fraction of muon candidates with nearby tracks grows as the cone opening angle increases. The agreement between data and MC is reasonably good. There are relatively few events with  $p_{T\Sigma track}^{\Delta R=0.X} > 0$  so most of the events are gathered in the first bin which is well reproduced by simulation. Figure 11.11 shows the relative isolation variable  $p_{T\Sigma track}^{\Delta R=0.X}/p_T$  and also these distributions show good agreement between simulation and data.

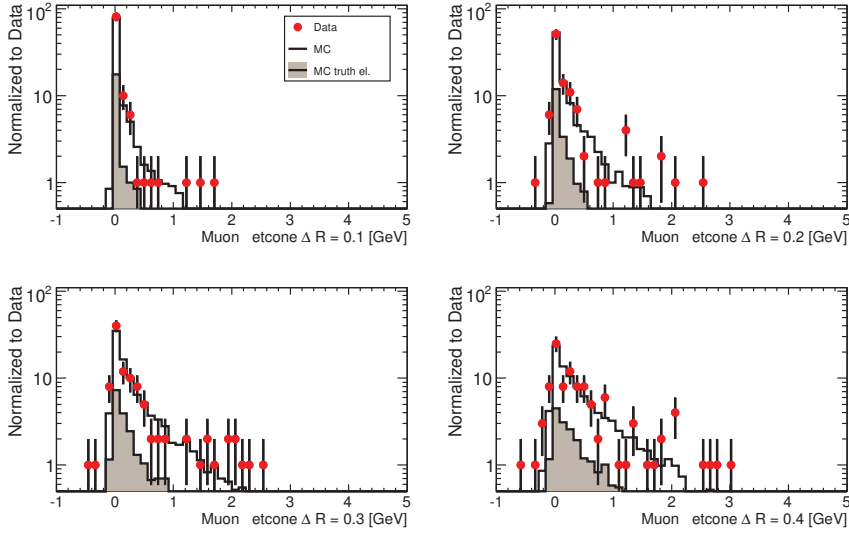


Figure 11.9: Comparison of data and MC of calorimeter-based isolation variable  $E_{\text{cal}}^{\Delta R=0.X}$  for  $\Delta R = 0.1, 0.2, 0.3, 0.4$  for all muons.

In several trilepton analyses, like for example in the CSC-notes [78], there is an isolation cut on the maximum- $p_T$  track in a cone around the muon track  $p_{T_{\text{max track}}}^{\Delta R=0.X}$ . The advantage of this variable is that it is less sensitive to soft tracks. It is by default not provided by the muon object and has to be calculated.

ATLAS provides an isolation calculation tool called TrackIsolationTool, however for muons, as applied in this analysis, it did not exclude the muon track itself. The  $p_{T_{\text{max track}}}^{\Delta R=0.X}$  was therefore calculated by a user defined tool. The input are the TrackParticleCandidate tracks (inner detector tracks) which pass the selection criteria. This means that only tracks of a certain quality are accepted, while the impact parameter cuts ensure that they originate from the primary vertex. The minimum track  $p_T$  is kept at 0.5 GeV as compared to 1.0 GeV in the default variables associated with the muon object. Figure 11.12 shows the  $p_{T_{\text{max track}}}^{\Delta R=0.X}$  for the four standard cone opening angles and there is a fairly good agreement between data and MC. Compared to the sum of track  $p_T$ ,  $p_{T\Sigma\text{ track}}^{\Delta R=0.X}$ , lowering the minimum  $p_T$  of the considered tracks shows that there are more muons with accompanying particles.

Also for the track-based isolation variables, the true muons represented by the shaded histograms are clearly most prominent in the first bin or at very low values.

## 11.5 Electrons

Electron candidates studied in this analysis are objects from the `ElectronAODCollection` container, the same as in the earlier SUSY searches presented in Chapter 8,9 and 10. Candidates have to pass the same selection cuts as muons, except for the muon specific algorithm requirement. Figure 11.13 shows that there is a good agreement between the data and MC distributions of the transverse impact parameter  $d_0$  (a) and the longitudinal

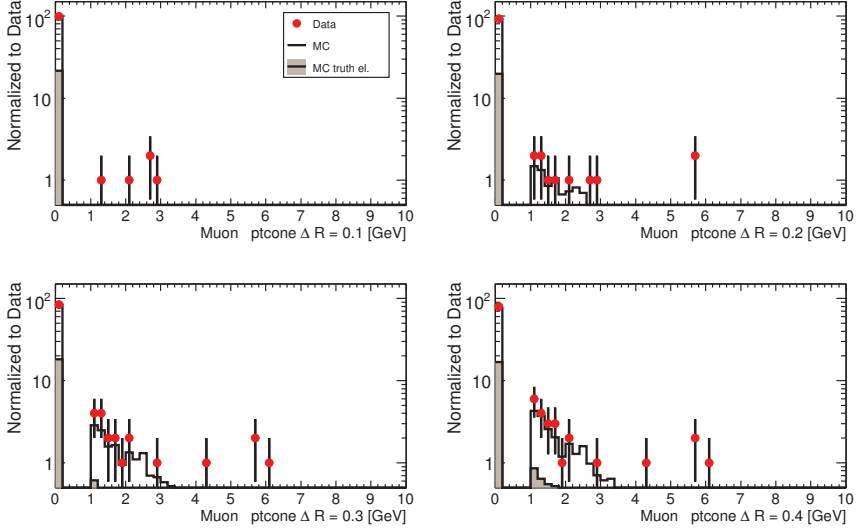


Figure 11.10: Comparison of data and MC muon track-based isolation. The isolation is a sum of track  $p_T$  in a cone. Minimum track  $p_T$  allowed in the sum is 1 GeV.

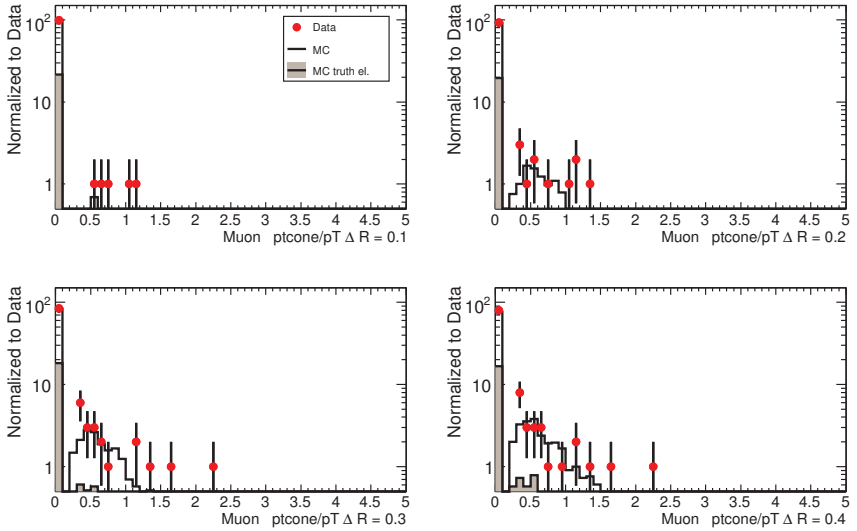


Figure 11.11: Comparison of data and MC muon relative track-based isolation. The isolation is the sum of track  $p_T$  in a cone divided by the muon  $p_T$ .

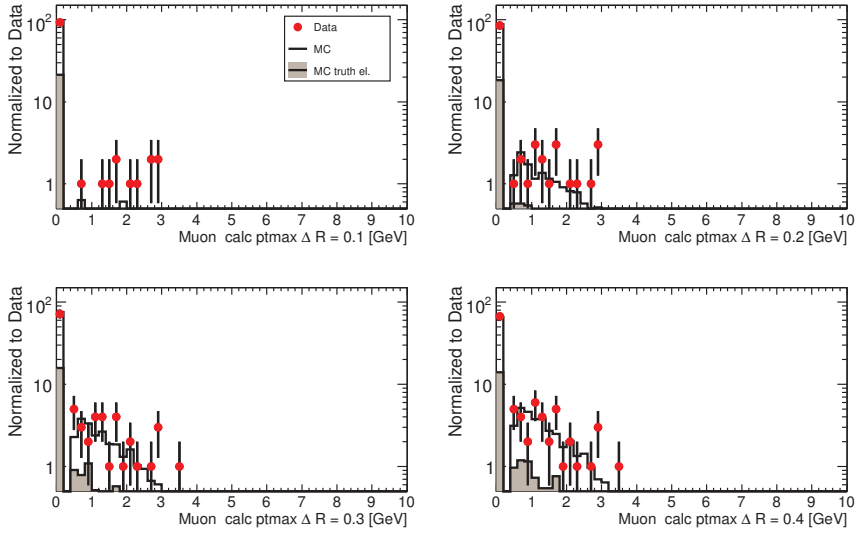


Figure 11.12: Comparison of data and MC muon track-based isolation, defined as the maximum  $p_T$  of tracks in a cone. Minimum track  $p_T$  is 500 MeV, to be compared with Figure 11.10

impact parameter  $z_0 \sin \theta$  (b). The gray shaded histogram, almost not visible, represents MC electron candidates matched to a MC truth electron within  $\Delta R = 0.03$ . Clearly the vast majority of the electron candidates are not real electrons, but rather, according to MC, pions, kaons or protons. The MC truth histogram will therefore be omitted from the next plots.

From Figure 11.14 showing the number of Pixel (a) and SCT (b) hits, we see that these too are well reproduced by the simulation.

A very important requirement in all analyses involving electrons is the electron reconstruction tag `isEM` (see Table 7.2). Figure 11.15 shows the distribution of the predefined masks tagging a loose (a), medium (b) and tight (c) electron. Histogram (c) focuses on the lower part of the histograms. The simulation is more optimistic giving larger fractions of loose, medium and tight electrons. Naturally enough, there are only very few tight electrons in the current sample. After applying the impact parameter and track quality cuts, approximately 98% of the electron candidates are in fact hadronic fakes, mainly pions and kaons. In order to keep as high statistics as possible, loose electrons are chosen for the isolation study.

In ATLAS there are two main electron reconstruction algorithms (see Section 7.3), each associated with an “author” tag:

- 1 for the high  $p_T$  `egamma` algorithm,
- 2 for soft electron algorithm
- 3 if the electron candidate was reconstructed by both.

In the 14 and 10 TeV analyses it was required that the electron was reconstructed by the high  $p_T$  algorithm or by both (author tag 1 or 3). However, as Figure 11.16 (a) shows, the

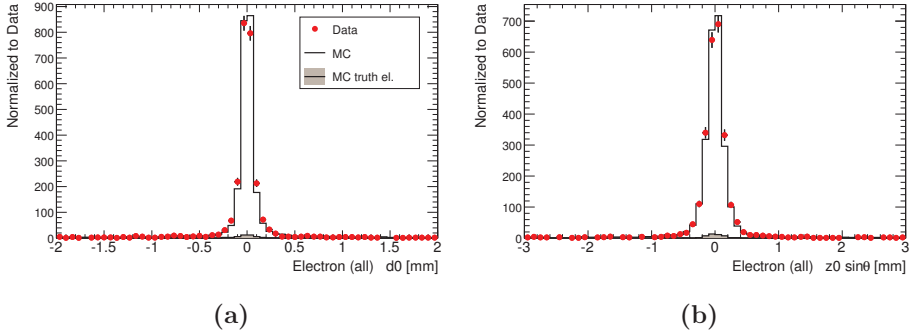


Figure 11.13: Comparison data-MC of the electron transverse impact parameter  $d_0$  (a) and the longitudinal impact parameter  $z_0 \sin \theta$  (b) measured with respect to the reconstructed primary vertex. No track quality cuts are applied.

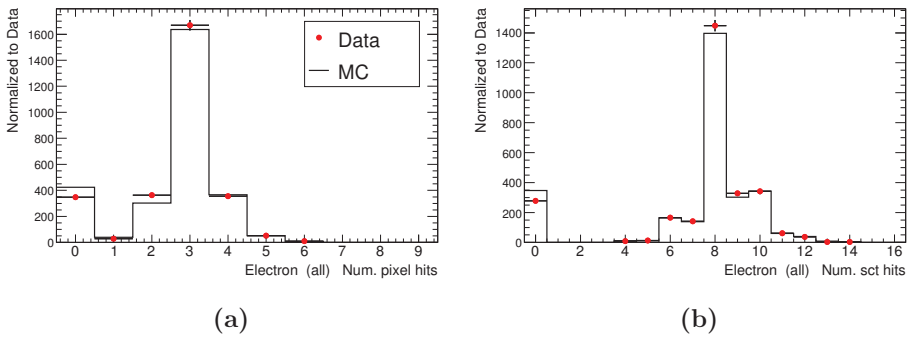


Figure 11.14: Comparison data-MC of the number of Pixel hits (a) SCT hits (b) for electrons. No impact parameter cuts are applied.

majority of the electron candidates in 900 GeV data are provided by the soft algorithm. In this plot the only requirement imposed was a minimum  $p_T$  of 500 MeV. This makes it possible to also include candidates provided by a forward electron algorithm represented by author 8. These are found in the very forward region of  $2.5 < |\eta| < 5.0$  with a  $p_T$  between  $\sim 5 - 15$  GeV. Even without any other requirements on these forward electron candidates, the  $p_T$  and  $\eta$  MC distributions match well with data (the distributions are not shown here as the forward electrons are not used in the analyses).

This analysis follows the common practice of rejecting electrons in the range  $1.37 < |\eta| < 1.52$ . Figure 11.16 (b) shows the calorimeter-based isolation in cone  $\Delta R = 0.3$  vs.  $\eta$ . Even before removing the crack  $\eta$  region, the distribution of isolation energy is even over the whole  $\eta$  range, possibly with some larger spread in the  $1.5 < |\eta| < 2.0$  range and around zero.

Figure 11.17 (a)-(b) show the comparison of the  $p_T$  for electron candidates provided by the soft and **egamma** algorithms, respectively. The soft algorithm is capable of providing electron candidates with  $p_T$  mainly in the range 1.5 - 3 GeV, while the **egamma** algorithm starts to be efficient at about 3 GeV.

Plots (c)-(d) in the same figure show the  $\eta$  distributions for the same two classes of electron candidates. Both the  $p_T$  and  $\eta$  distributions show reasonably good agreement



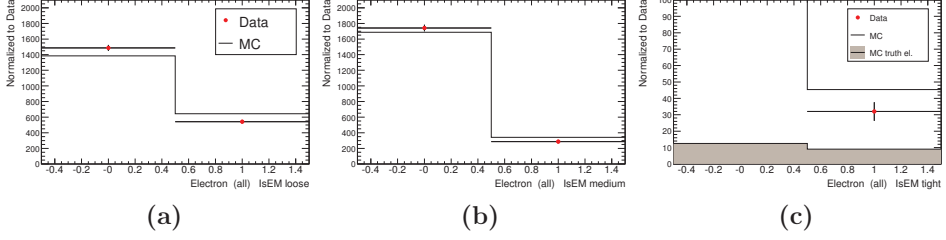


Figure 11.15: Comparison of electron  $\text{isEM}$  tags loose (a), medium (b) and tight (c). Histogram (c) shows only the lower part of the histogram. Inner detector hit and impact parameter cuts are applied.

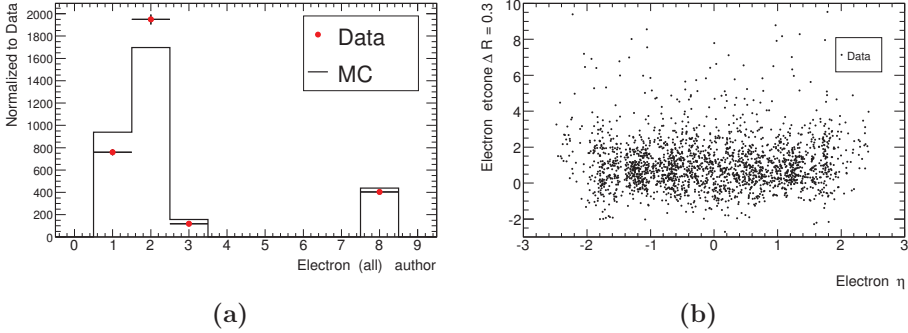


Figure 11.16: Distribution of  $e/\gamma$  and forward author tag for electron candidates in data and MC (a). In (a) there is no requirement on inner detector hits or impact parameter. Electron calorimeter-based isolation in cone  $\Delta R = 0.3$  vs.  $\eta$  (b). The impact parameter and track quality cuts are applied in (b).

between data and MC. The comparison is particularly interesting for the  $\eta$  distribution which has a very complex structure reflecting the structure of the ATLAS detector and subdetectors used in the reconstruction.

After all selection cuts and removal of the crack region there are 2028 electron candidates in the analyzed 900 GeV runs, 418 from the **egamma** algorithm and 1610 from the soft electron algorithm.

### 11.5.1 Electron isolation

Knowing that the vast majority of electron candidates are actually hadronic fakes, the main purpose of this analysis is to see how well MC reproduces the shapes of the various isolation variable distributions. The standard calorimeter-based isolation summed in a cone around the track is shown in Figure 11.18, (a) for candidates provided by the soft algorithm and (b) electrons passing the loose  $\text{isEM}$  criteria (no requirement on the author algorithm). There is a fairly good agreement between data and MC, however, for some of the cones the top of the MC distribution seems to be slightly shifted towards lower values compared to data. This discrepancy is more prominent in Figure 11.19 including all electrons candidates. The lower plots also show the difference between data and MC for cones with  $\Delta R = 0.2$  (a) and  $\Delta R = 0.4$  (b). The distributions of the isolation energy

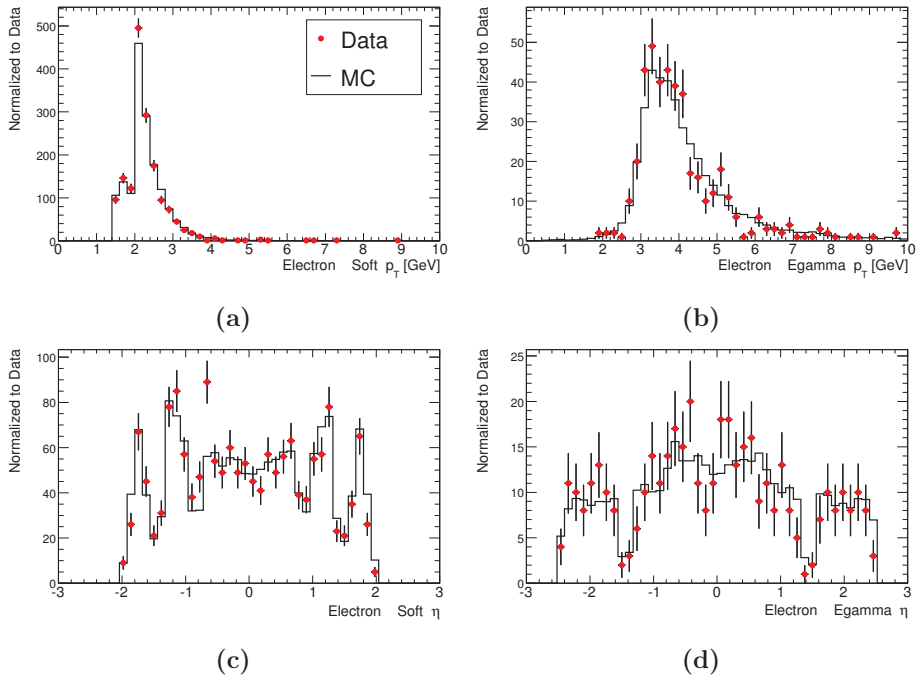


Figure 11.17: Comparison data-MC  $p_T$  for electrons from the soft algorithm (a) and `egamma` (b).  $\eta$  distributions for electrons from the soft algorithm (c) and `egamma` (d).

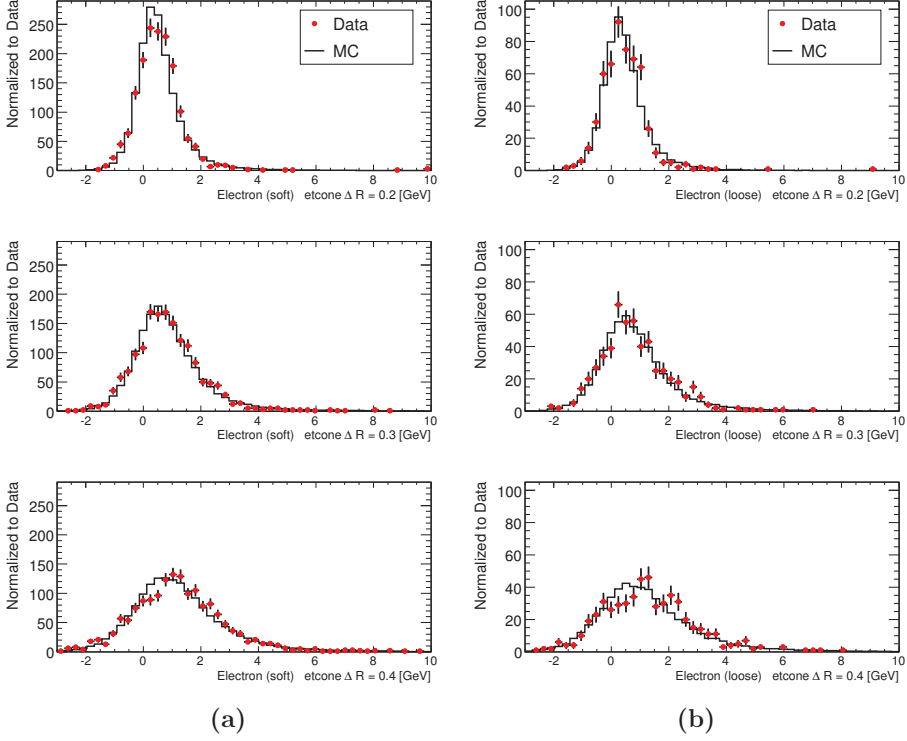


Figure 11.18: Comparison of data and MC electron calorimeter-based isolation  $E_{\text{cal}}^{\Delta R=0.X}$  for  $\Delta R = 0.2, 0.3, 0.4$ . Column (a) shows candidates from the soft algorithm, column (b) candidates which passed the **isEM** loose criteria (no author requirement).

in  $\Delta R = 0.3$ -cone is in better agreement as also indicated in the middle histogram row in Figure 11.18.

The electron object does not provide any track-based isolation. The sum of track  $p_T$  in a cone,  $p_{T\Sigma\text{track}}^{\Delta R=0.X}$ , has therefore to be calculated. In Figure 11.20 it is provided by the ATLAS TrackIsolationTool for cones with  $\Delta R = 0.1, 0.2, 0.3, 0.4$ , with the minimum  $p_T$  cut on tracks included in the sum of 1 GeV. The distributions include electron candidates that pass the loose **isEM** criteria and there is in general good agreement between data and MC.

The relative track-based isolation has shown good performance in several analyses (Chapter 8 and 10). Figure 11.21 shows the distributions of  $p_{T\Sigma\text{track}}^{\Delta R=0.X}$  from the TrackIsolationTool divided by the electron  $p_T$  (loose electrons). Again, there is a good agreement between data and MC.

Finally, Figure 11.22 shows the isolation which is the  $p_T$  of the hardest track in a cone around the electron,  $p_{T\text{max track}}^{\Delta R=0.X}$ . It is calculated using the TrackIsolationTool and the track  $p_T$  threshold of 1 GeV is clearly visible. In general, the simulation reproduces well the data.

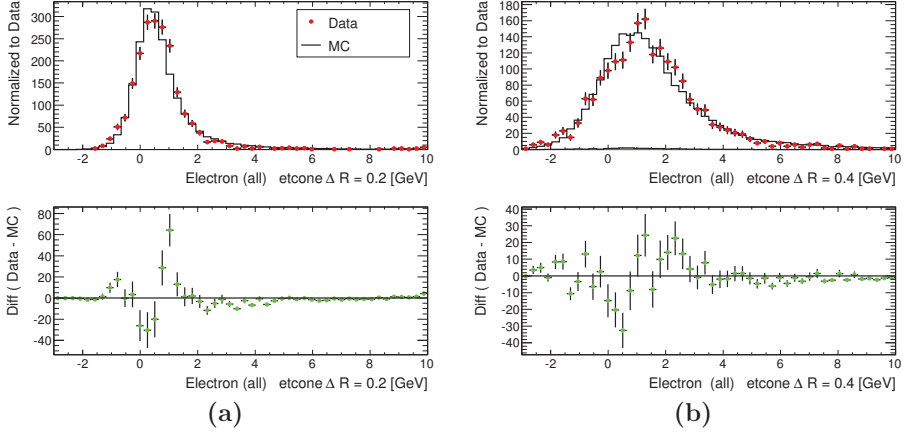


Figure 11.19: Comparison of data and MC electron calorimeter-based isolation  $E_{\text{cal}}^{\Delta R=0.X}$  for  $\Delta R = 0.2$  (a),  $\Delta R = 0.4$  (b) (all electron classes). The lower plots show the difference between data and MC.

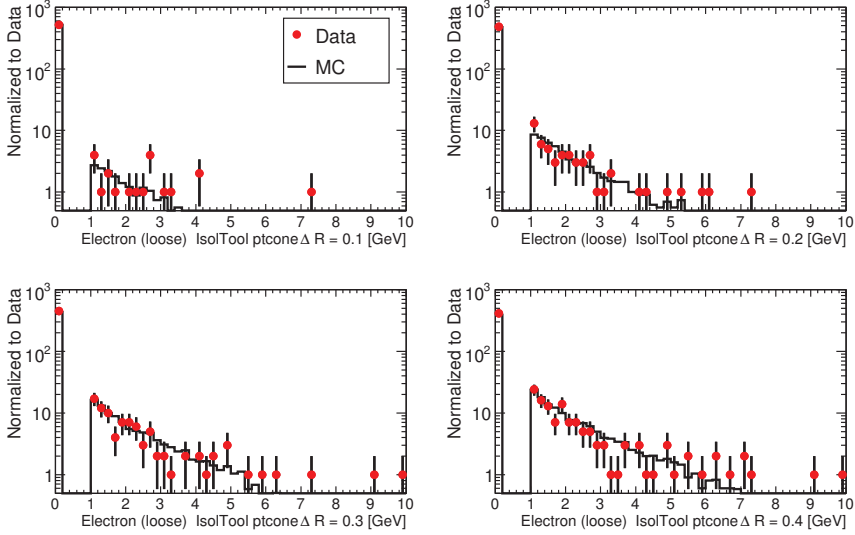


Figure 11.20: Comparison of data and MC track-based isolation  $p_{T\Sigma\text{track}}^{\Delta R=0.X}$  for loose electrons. The isolation is a sum of track  $p_T$  in a cone and it is calculated using an the TrackIsolationTool.

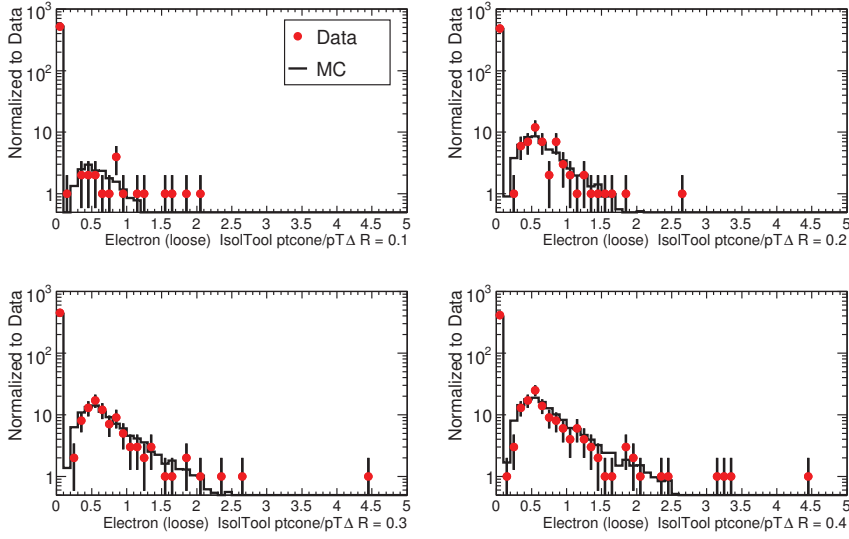


Figure 11.21: Comparison of data and MC relative track-based isolation  $p_{T\Sigma\text{ track}}^{\Delta R=0.X}/p_T$  for loose electrons. The isolation is a sum of track  $p_T$  in a cone from the TrackIsolationTool divided by the electron  $p_T$ .

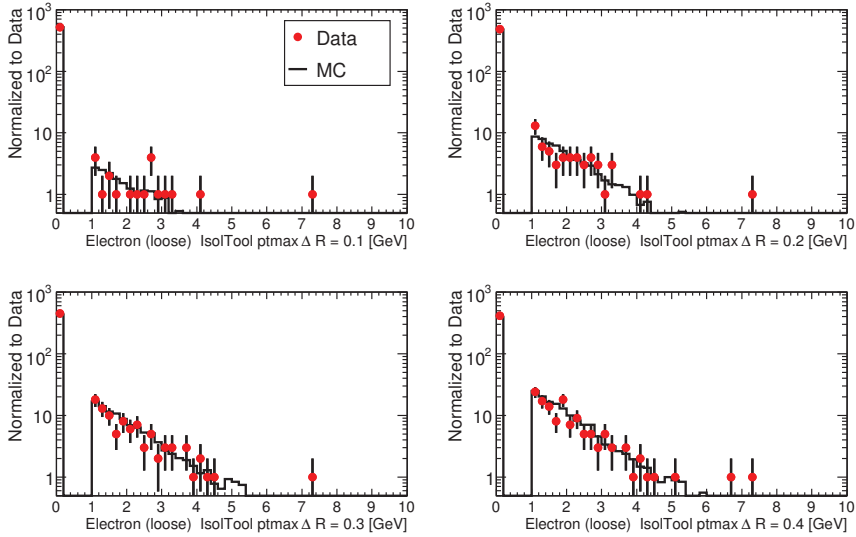


Figure 11.22: Comparison of data and MC track-based isolation  $p_{T\text{max track}}^{\Delta R=0.X}$  for loose electrons. The isolation is a the maximum  $p_T$  of tracks in a cone from the TrackIsolationTool.

### 11.5.2 Pions as electrons

One of the first particles that could be reconstructed after the 900 GeV collisions started was the  $K_S^0$ . It has a mass of 497.65 MeV and a branching ratio to  $\pi^+\pi^-$  of 62.20% [32].

In Figure 11.23 (a) which shows the invariant mass of tracks from a secondary vertex reconstructed from two tracks, assumed to be pions with opposite charge, one can clearly see a  $K_S^0$  peak. The vertical lines indicate the  $\pm 50$  MeV range around the nominal  $K_S^0$  mass. Tracks which lead to an invariant mass within this range,  $|m_{\pi^+\pi^-} - m_{K_S^0}| < 50$  MeV, were then matched to electron candidates. A match was accepted if the separation  $\Delta R$  was smaller than 0.3. With the current statistics there are only very few matches and in order to collect at least a handful of events no cuts other than  $p_T < 500$  MeV were applied.

Through this procedure one can obtain a sample of electron candidates which on the basis of a data-driven method can with high probability be considered as pions. The matching resulted in 29 candidates.

Figure 11.23 (b) shows the  $E/p$  distribution of the electron candidates matched to a pion track from  $K_S^0$ . The open histogram shows the MC distribution for electron candidates matched to a pion track from a  $K_S^0$ , while the gray shaded histogram represents MC candidates matched to a MC-truth electron. There is a clear difference between these two. Both MC histograms are normalized to data. For a real electron one expects the value of  $E/p$  to be around one (see the gray shaded histogram), while as shown by the data points, fake electrons will be shifted towards lower values.

Having a sample of pions one can start studying various features of fake electrons like shower shape variables and fake rates. With the low statistics of the current sample it is not feasible to do any precise analysis. However, it is still interesting to look at the isolation as it has been the focus of this study. Such fake electrons are expected to be less isolated as they are most probably part of a jet. Figure 11.24 shows the calorimeter-based isolation in cones with  $\Delta R = 0.2$  (a) and  $\Delta R = 0.3$  (b). Even with low statistics it is interesting to note the long tails as compared with the general distributions in Figure 11.18.

The last three figures show some of the shower shape variables used in the electron identification (see Table 7.2 and [78]). Figure 11.25 (a) shows the hadronic leakage defined as the ratio of the transverse energy deposited in a  $\Delta\eta \times \Delta\phi = 0.24 \times 0.24$  window in the

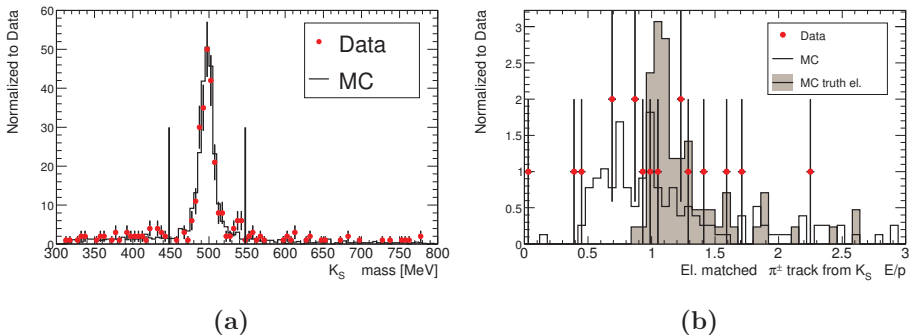


Figure 11.23:  $K_S^0$  mass (a), the vertical lines indicate the mass window where candidates were accepted as  $K_S^0$ .  $E/p$  (b) distribution for electron candidates matched to a track from a reconstructed  $K_S^0$ .

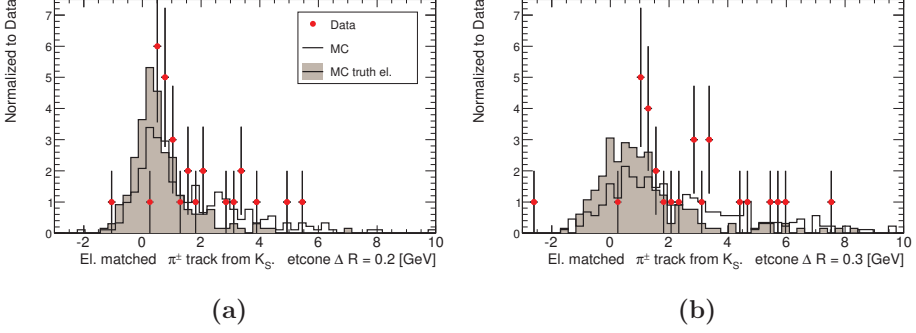


Figure 11.24: calorimeter-based isolation  $E_{\text{cal}}^{\Delta R=0.X}$  for electron candidates matched to a track from  $K_S^0$ , for  $\Delta R = 0.2$  (a),  $\Delta R = 0.3$  (b). No other cuts than  $p_T > 500$  MeV are applied.

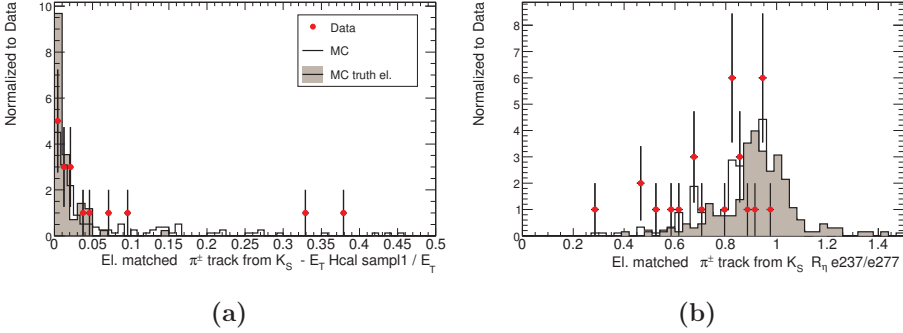


Figure 11.25: Hadronic leakage (a) defined as the ratio of the transverse energy deposited in a window  $\Delta\eta \times \Delta\phi = 0.24 \times 0.24$  in the first compartment of the hadronic calorimeter divided by the transverse energy deposited in the electromagnetic calorimeter. Lateral shower shape  $R_\eta(37)$  (b) defined as the ratio of energy reconstructed in  $\Delta\eta \times \Delta\phi = 3 \times 7$  cells to energy in  $\Delta\eta \times \Delta\phi = 7 \times 7$  cells. No other cuts than  $p_T > 500$  MeV are applied.

first compartment of the hadronic calorimeter divided by the transverse energy deposited in the electromagnetic calorimeter. Electrons are expected to be found close to zero while hadronic fakes in the tail at higher values.

Figure 11.25 (b) is the lateral shower shape  $R_\eta(37)$  given by the ratio of energy reconstructed in  $\Delta\eta \times \Delta\phi = 3 \times 7$  cells to energy in  $\Delta\eta \times \Delta\phi = 7 \times 7$  cells. As the electromagnetic shower is narrower than the hadronic one, electrons are expected to deposit most of their energy in a  $\Delta\eta \times \Delta\phi = 3 \times 7$  window in units of middle cells. The ratio  $R_\eta(37)$  is therefore close to one for electrons and distributed towards lower values for hadronic fakes.

The total shower width  $w_{\text{tot}}$  in unit of strip cells is shown in Figure 11.26. As expected, the electromagnetic showers are narrower than those associated with hadronic fakes.

Despite the low statistics, the distributions the electron candidates which are found to have a match to a pion, seem to follow the characteristic distributions expected for electrons faked by hadrons.

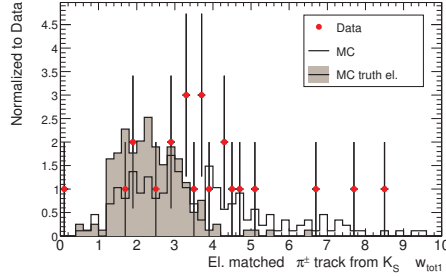


Figure 11.26: Total shower width  $w_{tot1}$  (in units of strip cells) in the first compartment of the electromagnetic calorimeter. No other cuts than  $p_T > 500$  MeV are applied.

## 11.6 Conclusions and outlook

The idea behind this brief study was to look at the physics objects used in the SUSY searches presented in the earlier chapters and compare the distributions of data and MC. Even at this early stage, there is good agreement.

Isolation which has a great importance in the trilepton searches has been of particular interest. A first look at the commonly used variables show in general good agreement both for muons and electrons.

The next step will be to study the lepton object definitions with data driven methods. This will involve for example tag-and-probe methods based on known signal like  $J/\psi$ ,  $Z$ . Top events which in many ways resemble SUSY events in some typical low mass scenarios may also be exploited.

A data driven method for study of pions faking electrons has been investigated. The low statistics of the 900 GeV samples allows only for a feasibility study, but it can be improved with higher statistics. All though ATLAS and the SUSY searches are optimized for high energies, it is very useful to understand lepton reconstruction also at low  $p_T$  as most fake leptons are actually soft.



# Chapter 12

## Systematic effects

The ATLAS experiment has during the last years been through enormous development in all fields. This thesis reflects some of the many changes. The earlier anticipated “early physics” scenarios changed from the nominal 14 TeV ( $1 \text{ fb}^{-1}$ ) to 10 TeV ( $200 \text{ pb}^{-1}$ ), for then finally to be set to 7 TeV for a two year data taking in 2010 and 2011 aiming at collecting  $1 \text{ fb}^{-1}$ . Although the definitions of the mSUGRA benchmark points stay the same, the phenomenological picture is changing. Due to the reduced center of mass energy, the focus has changed towards the low mass SUSY scenario represented in ATLAS by the SU4 benchmark point. The study has also been extended to a new set of points generated within the MSSM24 framework which represent a wider phenomenological picture than mSUGRA.

The software and computing framework have also constantly been changing and, comparing the object reconstruction efficiencies in Sections 7.5 and 9.2, the performance has been improving.

The chapter will start with some general comments (Section 12.1), then in Sections 12.2 and 12.3 discuss the systematic effects related to the 10 TeV and 14 TeV analysis in more detail.

### 12.1 Systematic effects - general comments

As all SUSY searches so far have been MC studies, the estimates of systematic uncertainties can only be vague. The realistic picture will only become clearer through studies of the detector performance, reconstruction and triggers based on real collision data. One of the most important tasks will be to gain better understanding of the SM background, luminosity, jet and  $\cancel{E}_T$  reconstruction. A full discussion of these challenging issues is beyond the scope of this thesis. However, based on the studies presented in previous chapters, the following section will point to some systematic effects.

#### 12.1.1 MC related effects

The limitations imposed by the lack of suitable MC samples play an important role. The comparison of the 14 TeV analyses and 10 TeV analyses which used a broader spectrum of SM background samples shows clearly that the  $Z$ +jets contribution is large for the trilepton channel, making a  $\cancel{E}_T$  cut, even a loose one, more important.

The MC statistics may also play an important role. For the 14 TeV analysis, the signal and background samples were relatively large. Unfortunately, the important  $Zb$  sample corresponded to significantly less than  $1 \text{ fb}^{-1}$ . In particular the jet exclusive analysis, where the distributions were normalized to  $10 \text{ fb}^{-1}$ , would profit from larger samples.

The main signal sample SU4 and SM backgrounds at 10 TeV are generated with integrated luminosities above  $200 \text{ pb}^{-1}$  to which they are normalized. The only exception is the QCD  $b\bar{b}$ +jets background for which the number of events correspond to very low luminosities leading to large uncertainty on the contribution from this process (see Table C.1 Appendix C).

Another important systematical effect is the currently limited understanding of the SM background. The MC simulation depends on a number of factors like Parton Distribution Functions (PDFs) and normalization. Simulation has to reproduce both the level and the shape of the distribution which also requires good understanding of the detector. Studies of the SM background will therefore be the first important phase of all searches for new phenomena.

Jet and  $E_T$  reconstruction are known to be very difficult tasks. During the early data taking one expects large uncertainties, especially related to tails of the distributions. Systematic effects due to jet energy scale and calibration will become better understood once they are studied with real data and these are not included in the study. However, awareness of these potentially large effects has motivated some of the analysis and cut strategies.

### 12.1.2 Significance measures

This search study uses the signal significance as the figure of merit. Four different definitions have been introduced in Section 8.1.1 and further discussed in Section 9.4. The results depend strongly on which definition one uses.

The 14 TeV analyses are optimized using the simple

$$S_0 = S/\sqrt{B} \quad (12.1)$$

which is known to be optimistic and it does not include uncertainties on background. Nevertheless, it is chosen as the main significance measure for the 14 TeV analysis in order to be more compatible with the CSC analyses.

At the same time, with  $1 \text{ fb}^{-1}$  the uncertainty related to the relevant backgrounds were estimated to be 20% [78] and, compared with the results given by

$$S_N = S/\sqrt{B + (\sigma_B)^2}, \quad (12.2)$$

it did not introduce drastic shift, as the background in most cases was relatively small. Depending on the background, the reduction in 14 TeV jet inclusive analysis was in the range 5-20%.

In order to be more realistic, the 10 TeV analyses used

$$Z_N = \sqrt{2} \operatorname{erf}^{-1}(1 - 2p) \quad (12.3)$$

described in Section 8.1.1 as the main significance measure. The advantage is that it treats the background uncertainty, which is expected to be high during the early analysis, in a more correct way. It is the most conservative significance measure.

In the early 10 TeV scenario ( $200 \text{ pb}^{-1}$ ) it is not unusual to set the background uncertainty to 100% or 50% [180]. Although 100% may be considered as a pessimistic estimate it has been chosen as the default value. However, the uncertainty will decrease as more data becomes available and the results are also shown for 50% background uncertainty.

The  $Z_N$  is also quoted for the 14 TeV analysis. However, it was calculated after the analysis was finalized using the event numbers presented in tables 8.4-8.8. Therefore the obtained values are presented with lower precision and they are not included in the discussion in this chapter.

As part of the study of significance measures, the 10 TeV analysis uses also the  $Z_{LLR}$ :

$$Z_{LLR} = \sqrt{2LLR} = \sqrt{2[(S+B)\ln(1+\frac{S}{B})-S]}. \quad (12.4)$$

During the cut optimization study it was observed that  $Z_{LLR}$  favours an event selection with mild cuts which leads to higher number of both signal and background events. In the analysis it was observed that for some cuts  $Z_{LLR}$  decreases as a result of lower number of events, while the other significances increase due to enhanced background suppression.

This may be related to the observations in Section 9.4. Figure 9.6 (b) shows that for small number of events  $Z_{LLR}$  tends towards the lower significance values given by  $Z_N$ . While it for large background, Figure 9.6 (c), is closer to  $S_0$ .

### 12.1.3 Trigger

A detailed study of the trigger efficiencies and systematical effect was not prioritized in this thesis. It is believed that the trilepton events will have a high trigger efficiency as the probability that all three leptons are below the trigger thresholds, which are typically 20-25 GeV, is very low. Section 8.3.2 estimates the efficiency of a logical OR-combination of a single electron and muon trigger to be above 90%.

A more detailed trigger study for trilepton events can be found in [78]. It shows that an OR-combination of the L2.e22i and L2.mu20 triggers give an efficiency of 93% for the SU2 direct gaugino events (SU2- $\chi\chi$ ) and 98% for SU3 events. These efficiencies are observed for events which have a same flavor opposite sign lepton pair and at least 3 isolated leptons.

## 12.2 10 TeV jet inclusive analysis

The 10 TeV analysis is the first one to be discussed as it includes a more orderly study of the systematic effects.

This section will first describe the study behind the choice of the jet-algorithm, muon-jet separation and an optimal set of event selection cuts. Section 12.2.2 will then describe a brief study of variation of the event selection cuts. Section 12.2.3 will summarize the results of the 10 TeV jet inclusive analysis and discuss some factors which are found to influence the SUSY trilepton discovery potential.

### 12.2.1 Cut optimization

In the 10 TeV analysis the focus is mainly on the low mass SU4 point. With a broader set of SM backgrounds than in the 14 TeV analysis, the strategy changed to a combination

of cuts on the  $p_T$  of the first two leading jets and  $\cancel{E}_T$ . The definition of the working point was done through a three dimensional cut scan performed for different combinations of jet algorithms (cone and Anti- $K_T$ ) and lepton-jet separations defined in Eq. 9.2-9.4. In addition, the analysis quotes the results obtained using the four significances mentioned in Section 12.1.2.

Table 12.1 shows results of the scan listing the sets of cuts giving highest significance for different combinations of jet algorithms and lepton-jet separation.

In most cases the quoted event selections correspond to the optimal cuts. However, for  $S_N$  with 100% background uncertainty, the selected set was also required to have moderate cut values, in particular on  $\cancel{E}_T$ . Moderate in this case means around 100 GeV. The resulting significances are somewhat lower, but avoid the situation where the results are based on a small number of events which is sensitive to the uncertainties related to the tails of the distributions and at the same time to the statistical fluctuations related to limited MC statistics.

It is also expected that during the early analysis there will be large systematical uncertainties related to the jet energy scale and  $\cancel{E}_T$  and it is an advantage to choose moderate cuts.

The scan shows that the highest significances are obtained using Anti- $K_T$  jets and the muon-jet separation given by Eq. 9.4 (see Table 12.1, bottom subtable). This separation requirement gives in fact the highest significances for both jet algorithms.

The set of event selection cuts defined as working point for the analysis was selected based on the highest significance as given by  $Z_N(\sigma_B/B = 100\%)$  and the cut values are:

$$p_T^{jet\ 1} > 80\ GeV, \quad p_T^{jet\ 2} > 50\ GeV, \quad \cancel{E}_T > 110\ GeV$$

The results of the scan are in agreement with the single lepton efficiency study (Sections 9.3.2, Table 9.5) which showed that this combination, Anti- $K_T$  jet and the muon-jet separation given by Eq. 9.4, gives the highest muon efficiency and the lowest fake rate.

The above listed cuts define the basis for the variation of the event selection cuts which will be presented in the next section.

### 12.2.2 Variation of cuts around the working point

The cuts were then varied within  $\pm 30$  GeV in steps of 10 GeV around the selected cuts, one at a time. Table 12.2 lists the results while Figure 12.1 shows a graphical presentation. In plot (b) and (c) the maximum values do not always correspond to the working point cuts. The discrepancy occurs because the working point is selected based on the optimization of  $Z_N$  while  $Z_{LLR}$  gives higher significances with looser cuts and  $S_N$  with tighter cuts, favouring low background.

The significances are least sensitive to variation of the cut on the leading jet which leads to a significance fluctuation of about 5%. The variation of the cuts on the  $p_T$  of the second leading jet and  $\cancel{E}_T$  leads to a fluctuation which roughly corresponds to a change in significance of 10%.

The results of this and previous section will now be presented in an overall summary of the 10 TeV jet inclusive analysis along with an overview of the systematic effects.

| Cone Jets                                       |                      |                      |                      |        |      |     |
|---|----------------------|----------------------|----------------------|--------|------|-----|
| $\Delta R(l, j) > 0.4$                          |                      |                      |                      |        |      |     |
| Cut:  | $p_T^{jet\ 1}$ [GeV] | $p_T^{jet\ 2}$ [GeV] | $\cancel{E}_T$ [GeV] | Signif | $S$  | $B$ |
| $S_0$   | 80                   | 50                   | 110                  | 28.5   | 41.8 | 2.2 |
| $Z_{LLR}$                                       | 80                   | 40                   | 50                   | 14.2   | 68.2 | 8.3 |
| $S_N(\sigma_B/B = 100\%)$                       | 90                   | 60                   | 110                  | 17.2   | 36.9 | 1.7 |
| $S_N(\sigma_B/B = 50\%)$                        | 90                   | 60                   | 110                  | 23.7   | 36.8 | 1.7 |
| $Z_N(\sigma_B/B = 100\%)$                       | 80                   | 60                   | 110                  | 8.4    | 37.8 | 1.8 |
| $Z_N(\sigma_B/B = 50\%)$                        | 80                   | 50                   | 100                  | 10.4   | 44.8 | 2.5 |
| $\Delta R > 0.4 \    \ p_T^l > 40 \text{ GeV}$  |                      |                      |                      |        |      |     |
| $S_0$   | 80                   | 50                   | 110                  | 31.7   | 50.2 | 2.5 |
| $Z_{LLR}$                                       | 60                   | 50                   | 50                   | 15.9   | 82.3 | 9.5 |
| $S_N(\sigma_B/B = 100\%)$                       | 100                  | 60                   | 110                  | 17.6   | 41.8 | 1.9 |
| $S_N(\sigma_B/B = 50\%)$                        | 90                   | 60                   | 110                  | 24.9   | 44.1 | 2.1 |
| $Z_N(\sigma_B/B = 100\%)$                       | 80                   | 50                   | 110                  | 9.0    | 50.2 | 2.5 |
| $Z_N(\sigma_B/B = 50\%)$                        | 80                   | 50                   | 110                  | 11.2   | 50.2 | 2.5 |
| $\Delta R > 0.4 \    \ p_T^{jet}/p_T^\mu < 0.5$ |                      |                      |                      |        |      |     |
| $S_0$   | 80                   | 50                   | 110                  | 32.4   | 52.0 | 2.6 |
| $Z_{LLR}$                                       | 60                   | 50                   | 50                   | 16.5   | 87.9 | 9.9 |
| $S_N(\sigma_B/B = 100\%)$                       | 110                  | 60                   | 110                  | 18.6   | 41.8 | 1.8 |
| $S_N(\sigma_B/B = 50\%)$                        | 80                   | 60                   | 110                  | 26.0   | 47.4 | 2.2 |
| $Z_N(\sigma_B/B = 100\%)$                       | 80                   | 60                   | 110                  | 9.2    | 47.4 | 2.2 |
| $Z_N(\sigma_B/B = 50\%)$                        | 80                   | 50                   | 110                  | 11.4   | 52.0 | 2.6 |
| Anti- $K_T$ -jets                               |                      |                      |                      |        |      |     |
| $\Delta R(l, j) > 0.4$                          |                      |                      |                      |        |      |     |
| Cut:  | $p_T^{jet\ 1}$ [GeV] | $p_T^{jet\ 2}$ [GeV] | $\cancel{E}_T$ [GeV] | Signif | $S$  | $B$ |
| $S_0$   | 80                   | 50                   | 100                  | 27.6   | 41.7 | 2.3 |
| $Z_{LLR}$                                       | 60                   | 50                   | 50                   | 13.7   | 64.0 | 8.0 |
| $S_N(\sigma_B/B = 100\%)$                       | 90                   | 60                   | 110                  | 16.5   | 33.0 | 1.6 |
| $S_N(\sigma_B/B = 50\%)$                        | 80                   | 50                   | 110                  | 22.7   | 38.5 | 2.0 |
| $Z_N(\sigma_B/B = 100\%)$                       | 80                   | 50                   | 110                  | 8.2    | 38.5 | 2.0 |
| $Z_N(\sigma_B/B = 50\%)$                        | 80                   | 50                   | 100                  | 10.12  | 41.7 | 2.3 |
| $\Delta R > 0.4 \    \ p_T^l > 40 \text{ GeV}$  |                      |                      |                      |        |      |     |
| $S_0$   | 80                   | 50                   | 110                  | 31.8   | 48.6 | 2.3 |
| $Z_{LLR}$                                       | 60                   | 50                   | 50                   | 15.5   | 79.5 | 9.5 |
| $S_N(\sigma_B/B = 100\%)$                       | 80                   | 60                   | 110                  | 17.5   | 44.3 | 2.1 |
| $S_N(\sigma_B/B = 50\%)$                        | 80                   | 50                   | 110                  | 25.3   | 48.6 | 2.3 |
| $Z_N(\sigma_B/B = 100\%)$                       | 80                   | 50                   | 110                  | 9.0    | 48.6 | 2.3 |
| $Z_N(\sigma_B/B = 50\%)$                        | 80                   | 50                   | 110                  | 11.17  | 48.6 | 2.3 |
| $\Delta R > 0.4 \    \ p_T^{jet}/p_T^\mu < 0.5$ |                      |                      |                      |        |      |     |
| $S_0$   | 80                   | 50                   | 110                  | 32.9   | 50.0 | 2.3 |
| $Z_{LLR}$                                       | 60                   | 50                   | 50                   | 16.0   | 83.7 | 9.8 |
| $S_N(\sigma_B/B = 100\%)$                       | 100                  | 70                   | 110                  | 18.3   | 35.6 | 1.5 |
| $S_N(\sigma_B/B = 50\%)$                        | 80                   | 50                   | 110                  | 26.1   | 50.0 | 2.3 |
| $Z_N(\sigma_B/B = 100\%)$                       | 80                   | 50                   | 110                  | 9.3    | 50.0 | 2.3 |
| $Z_N(\sigma_B/B = 50\%)$                        | 50                   | 50                   | 110                  | 11.4   | 52.8 | 2.6 |

Table 12.1: Jet inclusive cut optimization for SU4, 10 TeV analysis. Table shows the set of cuts giving highest significance.

| Anti- $K_T$ jets                                   |      |      |        |            |
|--|------|------|--------|------------|
| $\Delta R > 0.4 \parallel p_T^{jet}/p_T^\mu < 0.5$ |      |      |        |            |
|  | Min. | Max. | Fluct. | Fluct. [%] |
| $S_0 = 32.9$                                       |      |      |        |            |
| Cut $p_T^{jet\ 1}$                                 | 28.9 | 32.9 | 4.0    | -12.2      |
| Cut $p_T^{jet\ 2}$                                 | 26.7 | 32.9 | 6.2    | -18.9      |
| Cut $\cancel{E}_T$                                 | 28.1 | 32.9 | 4.8    | -14.6      |
| $Z_{LLR} = 15.0$                                   |      |      |        |            |
| Cut $p_T^{jet\ 1}$                                 | 13.5 | 15.3 | 1.8    | -12.0      |
| Cut $p_T^{jet\ 2}$                                 | 12.0 | 15.2 | 3.2    | -21.3      |
| Cut $\cancel{E}_T$                                 | 12.9 | 15.3 | 2.4    | -16.0      |
| $S_N(\sigma_B/B = 100\%) = 18.0$                   |      |      |        |            |
| Cut $p_T^{jet\ 1}$                                 | 16.6 | 18.0 | 1.4    | -7.8       |
| Cut $p_T^{jet\ 2}$                                 | 13.2 | 18.2 | 5.1    | -28.3      |
| Cut $\cancel{E}_T$                                 | 11.4 | 18.0 | 6.6    | -36.7      |
| $S_N(\sigma_B/B = 50\%) = 26.1$                    |      |      |        |            |
| Cut $p_T^{jet\ 1}$                                 | 23.5 | 26.1 | 2.6    | -10.0      |
| Cut $p_T^{jet\ 2}$                                 | 20.7 | 26.1 | 5.4    | -20.7      |
| Cut $\cancel{E}_T$                                 | 18.7 | 26.1 | 7.4    | -28.4      |
| $Z_N(\sigma_B/B = 100\%) = 9.3$                    |      |      |        |            |
| Cut $p_T^{jet\ 1}$                                 | 8.5  | 9.3  | 0.8    | -8.6       |
| Cut $p_T^{jet\ 2}$                                 | 7.7  | 9.3  | 1.6    | -17.2      |
| Cut $\cancel{E}_T$                                 | 7.6  | 9.3  | 1.7    | -18.3      |
| $Z_N(\sigma_B/B = 50\%) = 11.4$                    |      |      |        |            |
| Cut $p_T^{jet\ 1}$                                 | 10.4 | 11.4 | 1.0    | -8.8       |
| Cut $p_T^{jet\ 2}$                                 | 9.6  | 11.4 | 1.8    | -7.8       |
| Cut $\cancel{E}_T$                                 | 10.2 | 11.4 | 1.2    | -10.5      |

Table 12.2: Summary of variation of the three jet inclusive analysis cuts for SU4. The cuts were varied within  $\pm 30$  GeV in steps of 10 GeV around the working point ( $p_T^{jet\ 1} > 80$  GeV,  $p_T^{jet\ 2} > 50$  GeV,  $\cancel{E}_T > 110$  GeV), one at a time. The fluctuation *Fluct.* is  $(Sig_{max} - Sig_{min})$  where *Sig* denotes one of the significance measures, while the last column gives the fluctuation of the significance measured at the working point. A graphical presentation is shown in Figure 12.1.

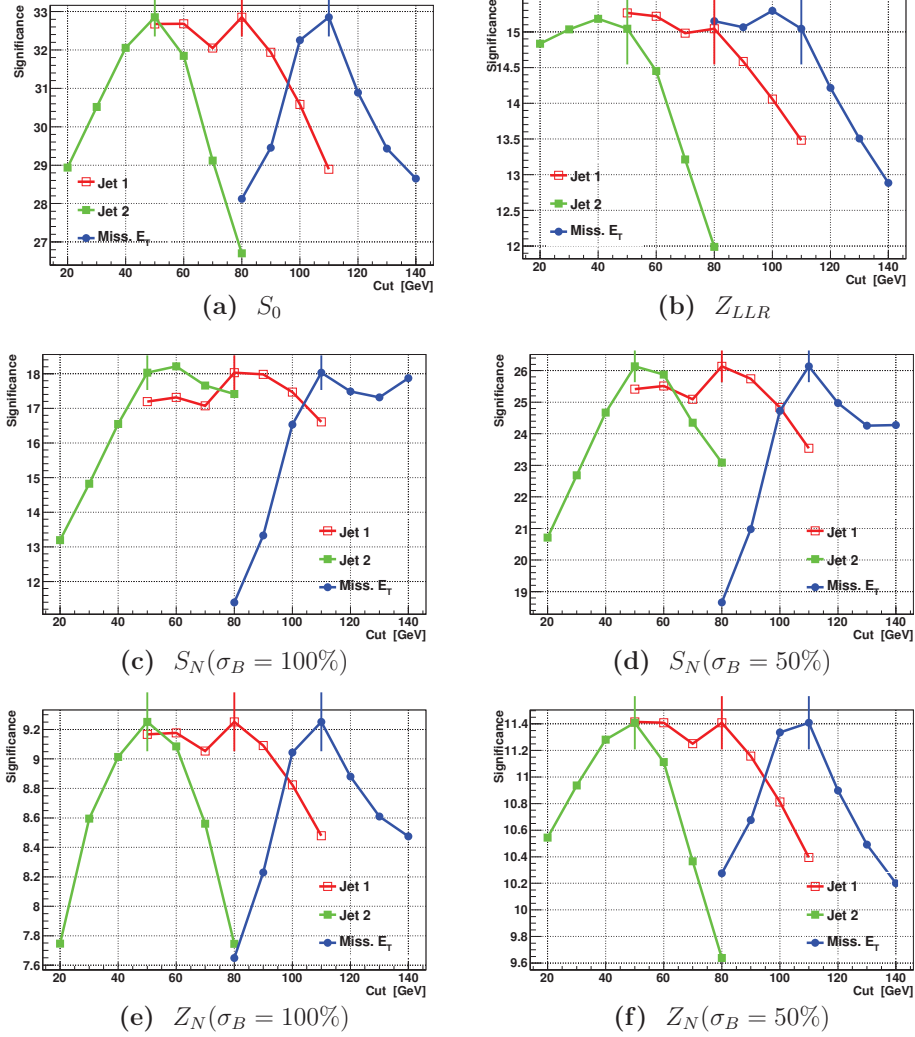


Figure 12.1: Graphical presentation of significances when varying the three jet inclusive analysis cuts for SU4. The cuts were varied within  $\pm 30$  GeV in steps of 10 GeV, one at a time. Note that the scales on the y-axis are different. The results are listed in Table 12.2.

### 12.2.3 Systematic effects: 10 TeV jet inclusive analysis

Motivated by the study presented in Section 12.2.1, the basis choice for 10 TeV is to use the Anti- $K_T$  jets and the muon-jet separation given by Eq. 9.4 (see Table 12.1, bottom subtable). The default isolation requirement is  $E_{\text{cal}}^{\Delta R=0.2} < 10$  GeV. It is a mild cut and not optimal for the trilepton analysis. However, the general recommendation for the early analysis was to build on the standards developed during the CSC analysis.

Table 12.3 summarizes, in the upper part, the results obtained with the baseline analysis, which is the reference point for the study of variations of the analysis presented in the text below.

This discussion will concentrate on the SU4 analysis presented in Section 9.5, but in most cases the observations are also valid for the other points, SU3 and SU1, which are summarized in Tables 12.4 and 12.5. These two points have low cross sections and due to the low number of signal events the systematic study can only be considered as indicative. This analysis has been optimized for SU4 and no study of lepton-jet separation or cut variation was performed for SU3 and SU1.

#### Lepton isolation

The 14 TeV analysis in Chapter 8 showed clearly that the trilepton analysis can be improved by an optimized isolation cut. The enumerated row 1. in Table 12.3 shows the results using an optimized calorimeter-based cut  $E_{\text{cal}}^{\Delta R=0.2} < 5$  GeV while row 2. shows the results for the relative track-based isolation  $p_{T\Sigma\text{track}}^{\Delta R=0.3}/p_T < 0.11$ .

With the exception of  $Z_{LLR}$ , which favours a working point with higher number of both signal and background events, the optimized calorimeter-based isolation cut improves the significances  $S_0$  and  $Z_N$  by about 5% and  $S_N$  by 20%. The optimized track-based isolation cut increases  $S_0$  by 13% and  $Z_N$  by 4%. Assuming 100% background uncertainty  $S_N$ , which is strongly enhanced by the reduced background, increases as much as 70%.

#### Jet algorithm

With the muon-jet separation given by Eq. 9.4 as default, the analysis was also performed using the Cone jets. The results can be found in the enumerated row 3. in Table 12.3. Even if the cut scan shows that a higher cut  $p_T^{\text{jet } 2} > 60$  GeV is slightly better when using cone jets, changing jet algorithm leads to only small significance reduction of about 1-5%. These results show that the event selection is not very sensitive to the choice of jet algorithm.

#### Event selection cuts

The fluctuation resulting from the variation of the three event selection cuts as presented in Table 12.2 has been added in quadrature. The sum can be found in the enumerated row 4 and is between 12 and 23%.

#### Lepton-jet separation

An estimate of the impact of the lepton jet separation is obtained from Table 12.1. The numbers represent the change when going from the best performing requirement (Eq. 9.4 Table 12.1, bottom subtable) to the  $\Delta R(l, j) > 0.4$  requirement which give lowest significances. The changes are about 10-15%.



### Background uncertainty

As one expects poor knowledge of the SM background during the early analysis, the uncertainty has been set to 100%. However, it is expected to be improved as more data become available. The significances were therefore also calculated with 50% background uncertainty. For SU4, this leads to an increase in  $Z_N$  and  $S_N$  of 20% and 45% respectively.

### SM background

Although this is a poor way of introducing background uncertainty, the analysis was repeated multiplying the SM background by a factor of two. The significance measures which do not include background uncertainty still reach  $5\sigma$  for SU4. However the significances are reduced by 30% ( $S_0$ ) and 11% ( $Z_{LLR}$ ). Admittedly, the method is very approximate, however it indicates that the signal to background ratio for SU4 is high enough to give a  $5\sigma$  discovery even if the background is doubled. The significance measures  $S_N$  and  $Z_N$  include the background fluctuations in a more correct way.

### NLO - LO

Some of the MC samples were generated using Next-to-Leading Order (NLO) generators, for the remaining ones NLO cross sections were obtained using K-factors from [78, 170]. NLO calculations represent a more correct theoretical picture and will be used in the analysis. However, there are also systematic uncertainties related to these calculations. These uncertainties were not available, so in order to obtain a very approximate estimate, the NLO cross sections were replaced by the Leading Order (LO) ones.

Typical values of K-factors are between 1 and 1.5. Comparing the results of the analysis with LO cross sections, the significances decrease by 5-15%. The only exception is  $S_N(\sigma_B/B = 100\%)$  which increases by about 7%. This can be explained by the fact that the K-factor for  $t\bar{t}$ , which is one of the main backgrounds, is larger than for the signal. Switching from NLO to LO the  $t\bar{t}$  background is more reduced than the SUSY signal and this causes the increase of  $S_N$ .

## 12. Systematic effects

|   |       |      |           |      |       |      |       |      |  |
|---|-------|------|-----------|------|-------|------|-------|------|--|
| Base line:  |       |      | SU4       |      |       |      |       |      |  |
| - Anti- $K_T$ jets, muon-jet separation Eq. 9.4, isolation $E_{\text{cal}}^{\Delta R=0.2} < 10$ GeV                     |       |      |           |      |       |      |       |      |  |
| - $p_T^{\text{jet } 1} > 80$ GeV, $p_T^{\text{jet } 2} > 50$ GeV, $\cancel{E}_T > 110$ GeV, 100% background uncertainty |       |      |           |      |       |      |       |      |  |
|   | $S_0$ |      | $Z_{LLR}$ |      | $S_N$ |      | $Z_N$ |      |  |
| S=50.0, B=2.3   | 32.8  |      | 16.6      |      | 18.0  |      | 9.2   |      |  |
| Change relative to the base line:   |       |      |           |      |       |      |       |      |  |
| 1. $E_{\text{cal}}^{\Delta R=0.2} < 5$ GeV  | 34.7  | +6%  | 16.1      | -3%  | 21.6  | +20% | 9.7   | +5%  |  |
| 2. $p_{T\Sigma\text{track}}^{\Delta R=0.3}/p_T < 0.11$  | 37.2  | +13% | 13.5      | -18% | 30.7  | +70% | 9.6   | +4%  |  |
| 3. Cone Jets  | 32.4  | -1%  | 16.8      | -1%  | 17.1  | -5%  | 9.1   | -1%  |  |
| 4. Event selection cut (*)  |       | -27% |           | -29% |       | -47% |       | -27% |  |
| 5. Lepton-jet separation (**)   |       | -16% |           | -14% |       | -10% |       | -11% |  |
| Background  |       |      |           |      |       |      |       |      |  |
| $\sigma_B/B$ 100% $\rightarrow$ 50%   |       |      |           |      | 26.1  | +45% | 11.4  | +25% |  |
| SM $\times$ 2   | 23.2  | -30% | 14.5      | -11% |       |      |       |      |  |
| NLO $\rightarrow$ LO  | 29.8  | -10% | 14.4      | -14% | 18.9  | +10% | 8.7   | -5%  |  |

Table 12.3: Systematic effects for SU4 jet inclusive analysis at 10 TeV. All numbers are for  $1 \text{ fb}^{-1}$ . Row 4 (\*) are from Table 12.2 fluctuation (rightmost column) added in quadrature. Row 5 (\*\*) are from Table 12.1 comparison with the  $\Delta R(l, j) > 0.4$  requirement which give lowest significances.

| Base line:  |       | SU3  |           |      |       |       |       |       |  |
|---|-------|------|-----------|------|-------|-------|-------|-------|--|
| - Anti- $K_T$ jets, muon-jet separation Eq. 9.4, isolation $E_{\text{cal}}^{\Delta R=0.2} < 10$ GeV                     |       |      |           |      |       |       |       |       |  |
| - $p_T^{\text{jet } 1} > 80$ GeV, $p_T^{\text{jet } 2} > 50$ GeV, $\cancel{E}_T > 110$ GeV, 100% background uncertainty |       |      |           |      |       |       |       |       |  |
|   | $S_0$ |      | $Z_{LLR}$ |      | $S_N$ |       | $Z_N$ |       |  |
| S=5.0, B=2.3  | 3.3   |      | 3.5       |      | 1.8   |       | 1.5   |       |  |
| Change relative to the base line:   |       |      |           |      |       |       |       |       |  |
| $E_{\text{cal}}^{\Delta R=0.2} < 5$ GeV   | 3.6   | +9%  | 3.4       | - 3% | 2.2   | + 20% | 1.7   | +14%  |  |
| $p_{T\Sigma\text{track}}^{\Delta R=0.3}/p_T < 0.11$   | 4.1   | +24% | 3.2       | - 9% | 3.4   | + 89% | 2.0   | +35%  |  |
| Cone Jets   | 3.3   |      | 3.5       |      | 1.7   | -6%   | 1.4   | -7%   |  |
| Background  |       |      |           |      |       |       |       |       |  |
| $\sigma_B/B$ 100% $\rightarrow$ 50%   |       |      |           |      | 2.6   | +45%  | 2.0   | + 33% |  |
| SM $\times$ 2   | 2.3   | -30% | 3.0       | -14% |       |       |       |       |  |
| NLO $\rightarrow$ LO  | 3.1   | -6%  | 3.1       | -11% | 2.0   | + 11% | 1.5   |       |  |

Table 12.4: Summary of systematic effects for the SU3 jet inclusive analysis at 10 TeV. All numbers are for  $1 \text{ fb}^{-1}$ .

|  |       |      |           |      |       |       |       |       |
|--|-------|------|-----------|------|-------|-------|-------|-------|
| Base line:   |       | SU1  |           |      |       |       |       |       |
| - Anti- $K_T$ jets, muon-jet separation Eq. 9.4, isolation $E_{\text{cal}}^{\Delta R=0.2} < 10$ GeV                      |       |      |           |      |       |       |       |       |
| - $p_T^{\text{jet } 1} > 80$ GeV, $p_T^{\text{jet } 2} > 50$ GeV, $\cancel{E}_T > 110$ GeV, 100% background uncertainty. |       |      |           |      |       |       |       |       |
|  | $S_0$ |      | $Z_{LLR}$ |      | $S_N$ |       | $Z_N$ |       |
| S=3.0, B=2.3   | 1.9   |      | 2.4       |      | 1.1   |       | 0.9   |       |
| Change relative to the base line:  |       |      |           |      |       |       |       |       |
| $E_{\text{cal}}^{\Delta R=0.2} < 5$ GeV  | 2.2   | +16% | 2.4       |      | 1.4   | + 30% | 1.1   | + 22% |
| $p_{T\Sigma\text{track}}^{\Delta R=0.3}/p_T < 0.11$  | 2.8   | +47% | 2.4       |      | 2.3   | +100% | 1.5   | + 67% |
| Cone Jets  | 1.9   |      | 2.4       |      | 1.0   | -10%  | 0.9   |       |
| Background   |       |      |           |      |       |       |       |       |
| $\sigma_B/B$ 100% $\rightarrow$ 50%  |       |      |           |      | 1.5   | +40%  | 1.2   | +37%  |
| SM $\times$ 2  | 1.4   | -26% | 2.1       | -13% |       |       |       |       |
| NLO $\rightarrow$ LO   | 1.8   | -5%  | 2.1       | -12% | 1.2   | +10%  | 0.9   |       |

Table 12.5: Summary of systematic effects for the SU1 jet inclusive analysis at 10 TeV. All numbers are for  $1 \text{ fb}^{-1}$ .

## 12.3 14 TeV analysis

The 14 TeV analysis with  $1 \text{ fb}^{-1}$  was the first early physics study in this thesis. The analysis is related to the CSC effort and follows some of the strategies developed in [78]. The main difference lies in the choice of lepton-jet separation and lepton isolation requirement. The original analysis presented in [181] did not include any methodical study of systematic effects. However, several strategies were developed, especially within the jet inclusive analysis, in order to prepare for various SUSY scenarios. These will be summarized here highlighting the effect of the variations. Section 12.3.1 will then give an overview of the observations.

The jet inclusive search at 14 TeV was performed along two paths, one based on a cut on the leading jet  $p_T$  (Section 8.2.1) and one on a cut on  $\cancel{E}_T$  (Section 8.2.2). The cuts were optimized for  $S_0 = S/\sqrt{B}$ . Relying only on the leading jet or only on  $\cancel{E}_T$  requires relatively hard cuts and may introduce considerable systematic uncertainties related to the jet energy scale and the tails of the distributions. It is therefore an advantage to have complementary analysis strategies.

The analysis was done with both a harsh cut at 200 GeV and a looser one at 150 GeV, for both variables. The  $\cancel{E}_T$ -based event selection was also done with and without removing events with a dilepton pair with invariant mass in the  $Z$ -mass window defined as  $|M_{l+l-} - M_Z| > 6$  GeV.

Table 12.6 summarizes the significances for the studied jet inclusive event selections. These were presented in more detail in Tables 8.2-8.7 which also included number of signal and background events. The quoted  $S_N$  and  $Z_N$  are calculated with 20% background uncertainty. This estimate is done for an integrated luminosity of  $1 \text{ fb}^{-1}$  [78] and assumes better understanding of the background than during the early 10 TeV analysis with  $200 \text{ pb}^{-1}$ . The 20% uncertainty is assumed for all backgrounds except QCD multijet events for which it is estimated to be 50% for  $1 \text{ fb}^{-1}$ . This background is, however, not

| Tight > 200 GeV |                |       | SU1   |       |       | SU2   |       |       | SU3   |       |       | SU4   |       |       |
|-----------------|----------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $p_T^{jet\ 1}$  | $\cancel{E}_T$ | $M_Z$ | $S_0$ | $S_N$ | $Z_N$ | $S_0$ | $S_N$ | $Z_N$ | $S_0$ | $S_N$ | $Z_N$ | $S_0$ | $S_N$ | $Z_N$ |
| ✓               |                |       | 22.1  | 21.0  | 10    | 7.7   | 7.3   | 5     | 63.0  | 60.0  | 21    | 156.2 | 148.7 | 38    |
|                 | ✓              |       | 17.7  | 16.6  | 9     | 4.4   | 4.1   | 3     | 38.9  | 36.4  | 15    | 78.5  | 73.5  | 24    |
|                 | ✓              | ✓     | 37.5  | 37.2  | 12    | 11.0  | 10.9  | 5     | 89.6  | 89.0  | 20    | 211.0 | 209.6 | 33    |
| Loose > 150 GeV |                |       | SU1   |       |       | SU2   |       |       | SU3   |       |       | SU4   |       |       |
| $p_T^{jet\ 1}$  | $\cancel{E}_T$ | $M_Z$ | $S_0$ | $S_N$ | $Z_N$ | $S_0$ | $S_N$ | $Z_N$ | $S_0$ | $S_N$ | $Z_N$ | $S_0$ | $S_N$ | $Z_N$ |
| ✓               |                |       | 17.5  | 15.8  | 9     | 6.6   | 6.0   | 4     | 49.4  | 44.4  | 19    | 190.7 | 171.5 | N/A   |
|                 | ✓              |       | 11.6  | 9.5   | 7     | 3.2   | 2.6   | 2     | 27.3  | 22.4  | 13    | 89.8  | 73.6  | 32    |
|                 | ✓              | ✓     | 12.6  | 11.6  | 7     | 4.1   | 3.8   | 3     | 32.2  | 29.7  | 9     | 114.9 | 105.9 | 33    |

Table 12.6: Variation of the jet inclusive event selection at 14 TeV.  $M_Z$  stands for  $|M_{l+l-} - M_Z| > 6$  GeV. The tick-marks indicate the event selection cuts. Background uncertainty is 20%. The  $Z_N$  calculation is based on the rounded numbers in Tables 8.2-8.7 and are therefore quoted with lower accuracy. All numbers are for  $1 \text{ fb}^{-1}$ . N/A means that  $Z_N$  is not calculable.

relevant for the trilepton analysis.

The “trilepton+ $\cancel{E}_T$ ” analysis with cuts  $\cancel{E}_T > 200$  GeV with  $|M_{l+l-} - M_Z| > 6$  GeV leads to the overall highest significances. However, due to the tight  $\cancel{E}_T$  the results are sensitive to statistical fluctuations. The number of background events is smaller than one ( $B = 0.3$ ) and the simple significances  $S_0$  and  $S_N$  are no longer appropriate.

Neglecting the results for this event selection we observe that the analysis based on a cut on the leading jet  $p_T$  gives the best results.

Table 12.7 in the next section will give a detailed comparison of the discovery potential of the various analyses.

### Lepton isolation

Both the jet inclusive and exclusive analyses have been done with an optimized set of lepton isolation requirements. Their impact within the jet inclusive analysis can be roughly estimated through a comparison with the CSC analysis which only differs by the choice of lepton isolation and lepton-jet separation [78]. While the number of signal events is comparable, the SM background is reduced by nearly 75%, leading to significances almost twice as large. The summary Table 12.7 quotes, in the bottom row of each sub-table, the decrease of significance which follows with the CSC-analysis isolation cuts. For SU1 which is not included in the analysis in [78] the decrease of 60% is a conservative estimate based on what is observed for the other points.

### SM background

The weakness of  $S_0$  due to negligence of the background uncertainty may to a limited extent be helped by calculating the significance with for example doubled SM background. From the definition of  $S_0$  one sees a reduction of 30%. This is however not a proper treatment of background uncertainty and the result does not correspond to including 100% background uncertainty in  $S_N$  and  $Z_N$  which would be significantly more reduced. However it may be noted that  $S_0$ , for all points except SU2, still reach more than  $5\sigma$  discovery with doubled SM background.

Another large systematic effect in the 14 TeV analysis is the limited set of SM backgrounds. The jet inclusive analysis at 10 TeV in Section 9.5 shows that after the trilepton requirement the  $Z$ +jets background is by far the largest. However, contrary to  $t\bar{t}$ , it can

be efficiently reduced by an  $\cancel{E}_T$  cut.

### 12.3.1 Overall summary jet inclusive analysis at 14 TeV

Table 12.7 summarizes the systematic effects studied for the jet inclusive analysis at 14 TeV.

In general we observe that the tight cuts are most suitable for SU1, SU2 and in particular SU3. For SU4 the discovery potential is higher for the loose cuts. SU4 is characterized by low masses which in turn give smaller mass differences and thus softer  $p_T$  and  $\cancel{E}_T$  spectra.

For jet inclusive analysis one observes that loosening the cuts on  $p_T^{jet\ 1}$  by 50 GeV, which corresponds to 25% of the baseline cut value, leads to a change of approximately 20-25%.

For the “trilepton+ $\cancel{E}_T$ ” analysis the baseline cut is  $\cancel{E}_T > 200$  GeV which gives the highest significance. The  $\cancel{E}_T > 200$  GeV with  $|M_{l+l-} - M_Z| > 6$  GeV is not taken into account in this discussion as it is subject to large statistical fluctuations.

A comparison of the results for the looser  $\cancel{E}_T$  cut at 150 GeV with the baseline  $\cancel{E}_T > 200$  GeV cut shows that the significances are reduced by 30-40%. The SU4  $S_0$  significance increases by 14% while  $S_N$  is unchanged. Adding the cut on  $M_Z$  window to the  $\cancel{E}_T$ -cut  $\cancel{E}_T > 150$  GeV increases again the significances. Comparing the  $\cancel{E}_T > 150$  GeV with  $|M_{l+l-} - M_Z| > 6$  GeV with the baseline cut  $\cancel{E}_T > 200$  GeV shows that despite the improvement due to the  $M_Z$  window cut, all points, except for SU4, obtain a lower significance.

Table 12.7 includes also the effect of doubling the SM background. The effect of going from NLO to LO is taken from the 10 TeV analysis as both analyses use the same K-Factors.

The last line labelled “Isolation” shows the reduction of significance for the “trilepton+jet” analysis with the tight cut  $p_T^{jet\ 1} > 200$  GeV compared to the CSC analysis which used a different lepton jet separation and isolation requirement.

### 12.3.2 14 TeV jet exclusive analysis

Within the mSUGRA framework, the jet exclusive analysis is specialized for the direct gaugino production which is most relevant in the Focus Point region represented by the SU2 point. It has a relatively low cross section compared to the other studied mSUGRA points and the distributions were normalized to  $10\text{ fb}^{-1}$ . This analysis is not considered as an early study. The main goal was to develop a general strategy and the analysis did not include a dedicated study of systematic effects. Some observations were however done and these will be briefly summarized.

Table 12.8 is an overview of the variation of jet exclusive analysis (see Section 8.3) related to the jet veto (no jet with  $p_T > 80$  GeV, JV) and a cut on the normalized impact parameter ( $IP_N < 6$ ) as the last cuts. The reason why the results for SU2 and SU2- $\chi\chi$  are equal after event selection with JV is that only direct gaugino production events -  $\chi\chi$  - pass the requirement.

The cut on the normalized impact parameter  $IP_N$  increases the significance ( $S_0$ ) by 6% when applied in combination with the JV. It appears also to be a good alternative to the JV in scenarios where there are additional sources of trilepton events, for example gluino pair production like in the case of SU2. In the analysis of direct gaugino production

|   | SU1   |      |       |      | SU2   |      |       |      |
|---|-------|------|-------|------|-------|------|-------|------|
| Cuts  | $S_0$ |      | $S_N$ |      | $S_0$ |      | $S_N$ |      |
| $p_T^{jet\ 1} > 200$ GeV                                  | 21.1  |      | 21.0  |      | 7.7   |      | 7.3   |      |
| $p_T^{jet\ 1} > 200 \rightarrow 150$ GeV                  | 17.5  | -20% | 15.8  | -25% | 6.6   | -15% | 6.0   | -18% |
| $\cancel{E}_T > 200$ GeV                                  | 17.7  |      | 16.6  |      | 4.4   |      | 4.1   |      |
| $\cancel{E}_T > 200 \rightarrow 150$ GeV                  | 11.6  | -34% | 9.5   | -43% | 3.2   | -27% | 2.6   | -37% |
| $\cancel{E}_T > 150$ GeV adding $M_Z$                     | 12.6  | +9%  | 11.6  | +22% | 4.1   | +28% | 3.8   | +46% |
| $\cancel{E}_T > 200 \rightarrow \cancel{E}_T > 150 + M_Z$ | 12.6  | -29% | 11.6  | -30% | 4.1   | -7%  | 3.8   | -7%  |
| General variation   |       |      |       |      |       |      |       |      |
| SM $\times$ 2   | -30%  |      |       |      | -30%  |      |       |      |
| NLO $\rightarrow$ LO (*)                                  | -15%  |      | -15%  |      | -15%  |      | -15%  |      |
| Isolation (**)  | -60%  |      | -60%  |      | -50%  |      | -57%  |      |

|   | SU3   |      |       |      | SU4   |      |       |      |
|---|-------|------|-------|------|-------|------|-------|------|
| Cuts  | $S_0$ |      | $S_N$ |      | $S_0$ |      | $S_N$ |      |
| $p_T^{jet\ 1} > 200$ GeV                                  | 63.0  |      | 60.0  |      | 156.2 |      | 148.7 |      |
| $p_T^{jet\ 1} > 200 \rightarrow 150$ GeV                  | 49.4  | -22% | 44.4  | -26% | 190.7 | +22% | 171.5 | +15% |
| $\cancel{E}_T > 200$ GeV                                  | 38.9  |      | 36.4  |      | 78.5  |      | 73.5  |      |
| $\cancel{E}_T > 200 \rightarrow 150$ GeV                  | 27.3  | -30% | 22.4  | -38% | 89.8  | +14% | 73.6  | 0%   |
| $\cancel{E}_T > 150$ GeV adding $M_Z$                     | 32.2  | +18% | 29.7  | +33% | 114.9 | +28% | 105.9 | +44% |
| $\cancel{E}_T > 200 \rightarrow \cancel{E}_T > 150 + M_Z$ | 32.2  | -17% | 29.7  | -18% | 114.9 | +46% | 105.9 | +44% |
| General variation   |       |      |       |      |       |      |       |      |
| SM $\times$ 2   | -30%  |      |       |      | -30%  |      |       |      |
| NLO $\rightarrow$ LO (*)                                  | -15%  |      | -15%  |      | -15%  |      | -15%  |      |
| Isolation (**)  | -57%  |      | -63%  |      | -42%  |      | -50%  |      |

Table 12.7: Overall summary of studied systematical effects for jet inclusive analysis at 14 TeV. (\*) NLO $\rightarrow$ LO change is from the 10 TeV analysis. (\*\*) The effect of the different isolation requirements is obtained by comparing the results of the trilepton plus  $p_T^{jet\ 1} > 200$  GeV analysis with the corresponding results in the CSC analysis [78].

| Cuts |                 | SU2   |       |       | SU2- $\chi\chi$ |       |       |
|------|-----------------|-------|-------|-------|-----------------|-------|-------|
| JV   | IP <sub>N</sub> | $S_0$ | $S_N$ | $Z_N$ | $S_0$           | $S_N$ | $Z_N$ |
| ✓    |                 | 4.9   | 1.3   | 1     | 4.9             | 1.3   | 1     |
| ✓    | ✓               | 5.2   | 1.6   | 2     | 5.2             | 1.6   | 2     |
|      | ✓               | 10.5  | 2.5   | 3     | 5.5             | 1.5   | 3     |

Table 12.8: Variation of the jet exclusive event selection at 14 TeV. All numbers are for 1 fb<sup>-1</sup>.

scenarios generated within the MSSM24 framework presented in Section 10.1.1, the jet veto proved to be a good strategy.

The lepton isolation is also important in jet exclusive analysis, however, the main SM background, especially when applying the jet veto, is no longer  $t\bar{t}$  but the  $WZ$  production. However, the 14 TeV analysis is affected by the limited set of SM backgrounds. The analysis of the MSSM24 direct gaugino points at 10 TeV with a broader set of SM showed that the dominant contributions is from  $Z$ +jets events. The  $\cancel{E}_T$  cut, even a loose one, becomes very important as it efficiently reduces this background.

## 12.4 Conclusions

As all SUSY searches in ATLAS so far are based on MC, it is only possible to give vague estimates of systematic effects. This chapter investigated a number of effects influencing the trilepton SUSY searches, both jet inclusive and exclusive.

This chapter has summarized various systematic effects relevant to the 10 TeV and 14 TeV analysis. At this point one of the main sources of uncertainties is the fact that all SUSY searches are based on MC simulations. There are theoretical uncertainties related to both the level and the shape of the distributions. However, the NLO calculations are always being improved and study of the new LHC data will contribute to better tuning of the theoretical framework, for example improving the PDFs.

In addition, the MC-samples may have limited statistics or some processes may simply be unavailable. The extended set of backgrounds used in the 10 TeV analysis showed a different picture compared to the 14 TeV analysis.

However most of these systematic uncertainties related to MC and SM background are reducible. As larger collision data samples become available, one will get a more precise knowledge of the backgrounds through data-driven methods. The high background uncertainty used in the 10 TeV early scenario will surely be reduced. Chapter 11 presented a comparison of data and MC for the first  $\sqrt{s} = 900$  data. Already at this early stage there is a fairly good agreement between data and simulation indicating that the ATLAS performance is close to what is expected.

The nature of SUSY is not known and it is therefore important at this point to develop different search strategies. The 14 TeV jet inclusive analyses clearly showed that looser jet  $p_T$  and  $\cancel{E}_T$  cuts are favourable for the low mass scenarios, like SU4, while harder cuts give better results for the other points. It may also be an advantage to explore the performance of event selection strategies which uses different variables. The 14 TeV jet inclusive analysis followed the standard paths: cutting on the jet  $p_T$  and  $\cancel{E}_T$ . However, it is interesting to investigate other possibilities which may be less focused on mSUGRA characteristics. The study of the MSSM24 direct gaugino points showed that the non-

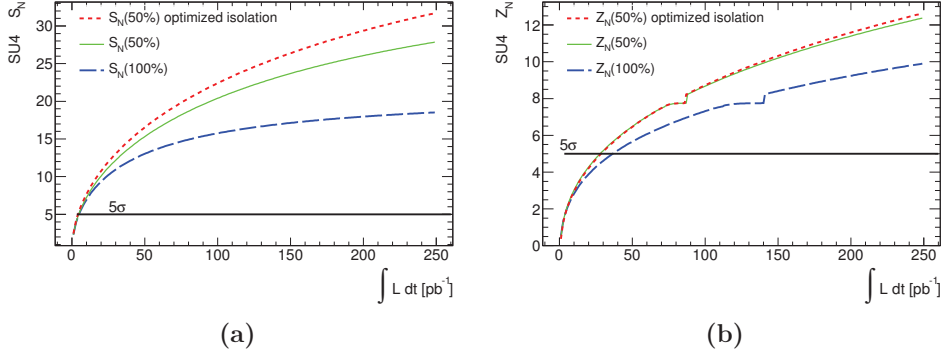


Figure 12.2: SU4  $S_N$  (a) and  $Z_N$  (b) plotted as function of the integrated luminosity for the jet inclusive baseline analysis at 10 TeV, with 100% and 50% background uncertainty, and for the optimized isolation and 50% background uncertainty. The number in parenthesis stands for background uncertainty.

typical mSUGRA jet exclusive analysis is applicable to a broad range of SUSY scenarios.

The 10 TeV analysis investigated also the effect of using the Anti- $K_T$  jet reconstruction algorithm instead of the fixed cone algorithm. With the choice of muon-jet separation as described in the analysis, changing the jet algorithm has only minor impact on the discovery potential.

Both the jet inclusive and to a certain extent, the jet exclusive analysis relies on jet  $p_T$  and  $\cancel{E}_T$ . During the early analysis it is expected that there will be large systematic uncertainties related to these quantities. However, this is also an effect that is expected to be reduced as the experiment collects more data.

The lepton isolation and lepton-jet separation has been an important topic throughout all analyses presented in this thesis. This summary highlights the improvements which can be gained by a good choice of variables and optimization of cuts.

### Summary systematic effects on SU4 -jet inclusive analysis at 10 TeV

In order to get an idea of the systematic effect related to the variation of the 10 TeV jet inclusive analysis, the relative changes listed in row 3-5 of the enumerated part of Table 12.3 were added in quadrature. The effects of systematics on the significances of the baseline analysis, for 200  $\text{pb}^{-1}$  and 100% background uncertainty, are quoted below:

- $S_0 = 32.8 \rightarrow 22.6$  (-31%)
- $Z_{LLR} = 16.6 \rightarrow 11.3$  (-32%)
- $S_N = 18.0 \rightarrow 9.3$  (-47%)
- $Z_N = 9.2 \rightarrow 6.7$  (-27%)

Figure 12.2 shows the significance  $S_N$  (a) and  $Z_N$  (b) as a function of the integrated luminosity. It is plotted for the baseline analysis with 100% and 50% background uncertainty. Then the last graph (red/dotted) represents the analysis using 50% background uncertainty and the track-based isolation for  $S_N$  and calorimeter-based for  $Z_N$ . This plot



shows the combined effect of these two important factors that improve the discovery potential. The lower background uncertainty gives higher significances for both measures, while the improved isolation has a larger effect on  $S_N$ .

$S_N$  reaches  $5\sigma$  for all three choices after about  $4\text{--}5\text{ pb}^{-1}$ . With higher integrated luminosity there is, however, a clear difference between the three options where track-based isolation with 50% background uncertainty gives clearly the highest significance. For the baseline analysis  $Z_N$  reaches  $5\sigma$  with  $38\text{ pb}^{-1}$ , while the luminosity required for  $5\sigma$  with 50% uncertainty, with or without optimized isolation, is  $28\text{--}29\text{ pb}^{-1}$ .

The calculation of  $Z_N$  is a numerical process which involves both a Poisson and Gaussian probability density function. The observed jump may be related to some numerical effects or approximation.

For the other studied benchmark points, SU1 and SU3, a  $5\sigma$  discovery lies in the  $\text{fb}^{-1}$  range with a background uncertainty at the level of 20%.

### Summary systematic effects on jet inclusive analysis at 14 TeV

The 14 TeV jet inclusive analysis presented in Chapter 8 shows that all studied benchmark points reach a  $5\sigma$  discovery within about  $1\text{ fb}^{-1}$ . For the jet exclusive analysis specialized for direct gaugino production, a  $5\sigma$  discovery lies at the  $10\text{ fb}^{-1}$  level.

It is more difficult to give an overall systematic effect for the 14 TeV jet inclusive analysis which is summarized in Table 12.7. However, it is observed that even with the relatively large variations of the event selection cuts, the significances change by no more than 10%–45%. Comparing the values of  $S_0$  and  $S_N$  one observes that the effect of the background uncertainty which is assumed to be 20% is relatively small as the number of background events is in most cases small.

$Z_N$  indicates clearly lower significances. However, it is interesting to note that with the exception of SU2, all other points reach  $5\sigma$  for all combinations of event selection cuts (Table 12.6).



# Summary and outlook

The high energy physics experiments have over the last decades strengthened the position of the Standard Model (SM). At the same time, its limitations clearly call for extended models among which Supersymmetry is considered as one of the most promising. In order to address these questions, CERN has designed and built the LHC accelerator and four large experiments. ATLAS is one of the two general purpose detectors specialized in study of proton collisions. The very ambitious design luminosity of  $10^{34}\text{cm}^{-2}\text{s}^{-1}$  and center of mass energy of 14 TeV require a very complex and large facility which has been very challenging to bring into full scale operation. There have been several delays of the start-up followed by revisions of the early physics scenarios which have influenced the structure of this thesis.

The LHC experiments are also facing enormous challenges in the area of computing. In order to make the data available to all members of the collaborations and share the work load, CERN has developed a global distributed Grid computing environment. NorduGrid is one of the three grid flavors of the Worldwide LHC Computing Grid. It connects the resources using the Advanced Resource Connector - ARC - middleware which organizes the resources as a truly distributed Tier-1. The ARC enabled Tier-1 requires its own adaptations of some of the components of ATLAS computing model.

This thesis reports on efforts in the area of Grid job monitoring and software distribution which were required at some point, but which became obsolete as central solution became available. It has also covered two major projects: managed Monte Carlo production and distributed analysis in the NDGF Tier-1. Both are currently fully integrated in the ATLAS computing environment and show excellent performance.

As the Grid managed production and analysis frameworks are indispensable tools, the contributions to the development and operation of these computing infrastructures are directly related to the main topic of the thesis: searches for Supersymmetry in trilepton final states. Such events are very interesting due to their clear leptonic signature and potentially low background. The requirement of at least three isolated leptons reduces strongly the expected large QCD background making searches for trilepton events interesting even during the early data taking.

At the center of mass energy of 10 TeV, if SUSY follows a low mass scenario as represented by the mSUGRA SU4 point, the jet inclusive search can yield a  $5\sigma$  discovery potential with roughly  $40\text{ pb}^{-1}$ . This estimate is based on the most conservative significance measure  $Z_N(\sigma_B/B = 100\%)$ . The other ATLAS SUSY benchmark points require an integrated luminosity in the  $\text{fb}^{-1}$  range, when the background uncertainty can be estimated at the level of 20% as in the 14 TeV analysis.

The 14 TeV jet inclusive analysis shows that a significant signal can be obtained for all benchmark points, except SU2, with less than  $1\text{ fb}^{-1}$ .

However, if SUSY follows a Focus point like scenario, the strongly interacting super-

symmetric particles become heavy and may be beyond the reach of LHC. Gauginos are then expected to be the lightest supersymmetric particles (heavier than the LSP which is  $\tilde{\chi}_1^0$ ) and direct production of gaugino pairs may be the only possible SUSY production process. In Focus point region also the gluino is typically light enough to be produced. Such scenario was investigated in the jet exclusive analysis at 14 TeV and the significance  $S/\sqrt{B}$  reaches  $5\sigma$  with approximately  $10 \text{ fb}^{-1}$ .

The main study has been based on the mSUGRA model which represents many phenomenologies considered as typical for SUSY. However, the anticipated unification at the GUT scale imposes certain mass hierarchies and phenomenological characteristics like a large number of high- $p_T$  jets and large  $\cancel{E}_T$ , which might not be the case for less constrained SUSY models.

The search strategies developed for mSUGRA were therefore applied to a new set of SUSY benchmark scenarios generated within the MSSM24 model. The jet exclusive analysis appeared to be suitable for a range of points with heavy sfermions and large cross sections for direct gaugino production. The jet inclusive search developed for the SU4 low mass point showed very good results for similar jet,  $\cancel{E}_T$  and lepton rich scenarios.

With the lack of knowledge of the ATLAS performance derived from real data it is natural to expect large systematic effects related to these results. The studies of systematic effects show that the background uncertainty is one of the most significant factor. The discovery potential of the MSSM24 points presented in Chapter 10 was calculated assuming 100% background uncertainty. With a background uncertainty of 50% one can expect that  $S_N$  improves by about 35% while  $Z_N$  by about 25%.

However, the situation is changing. ATLAS collected its first collision data in December 2009. The data-MC comparison study in Chapter 11 shows that even at this early stage, there is already a reasonably good agreement between MC and data which strengthens the credibility of the MC based analyses.

As the 7 TeV data becomes available, studies of the SM background and derivation of efficiencies and uncertainties through data driven methods will be the first priority. The coming year will be devoted to rediscovering and understanding the SM before one can hope to see any sign of new physics.

During the early physics program at 7 TeV, ATLAS SUSY searches will concentrate on low mass scenarios such as SU4. The cross section of this model is reduced to 47.9 pb which is roughly half the value of the cross section at 10 TeV.

Although it might not be possible to reach any significant excess of trilepton events over the SM background, with the planned  $1 \text{ fb}^{-1}$  of 7 TeV data to be collected during 2010 and 2011, ATLAS should soon be able to push the exclusion limits currently set by the Tevatron experiments or even see excesses of interesting events.

The last part of the thesis reported on outreach activities. With the start-up of the LHC the interest in particle physics and CERN has grown enormously. The exciting perspectives of new physics have caught many people's attention. It will of course be great to announce discoveries. The challenging part is, however, to explain that these may take time and that failures and repair are an expected part of the experimental work. We also have to prepare for a situation where little new is discovered and be able to communicate that whatever knowledge LHC reveals, it is of great value and extends our understanding of Nature.

# Part IV

## The Outreach Challenge



# Chapter 13

## Outreach activities

The high energy physics experiments have grown to spectacular dimensions and the questions asked have stretched people's imagination beyond what one could expect in science-fiction. Many people in the general public have begun realizing how fundamental the knowledge of particle physics is for our understanding of the universe and its development. The progress of this academic adventure depends strongly on support from the societies, not only in terms of interest, but also through highly concrete and sustainable economic grants. It is therefore a highly justified expectation that scientific communities disseminate the knowledge to a broader public.

Norway, as many other countries in Europe, has for many years experienced a decline of students choosing the field of physics [182] [183]. In the longer run, a continuous recruitment of good candidates is essential for building the future research communities. Society in general has an escalating need for persons educated in the field of natural science, particularly physics and mathematics, to keep up with technological development and the solving of energy and climate related problems [184]. Outreach and education activities are therefore, for many important reasons, a good investment in the future.

This chapter will briefly report on the main outreach efforts. Personally, I have found outreach endeavors very interesting and inspiring and thus have added it to the initial project description as an integral part of the PhD work. It is noted that the project plan was accepted at admission to the PhD program.

### 13.1 Masterclass 2005-2008

The year 2005 was declared by the World Congress of Physical Societies to be the World Year of Physics [185] and at the same time an Einstein jubilee. The European Particle Physics Outreach Group (EPPOG) [186] wanted to celebrate it by organizing an international event for high school students (17-18 years old). A suitable one day program called "Masterclass in Particle Physics" was already well established in England and was chosen as a template for the international project.

After a very positive experience in 2005, EPPOG decided to continue organizing further international events. The author of this thesis was appointed to take the lead in developing and organizing the Masterclasses in Oslo in 2005 - 2008.

The main ideas of Masterclass are:

- Invite high school students to a real research environment

- Offer high quality lectures by people directly involved in the field
- Give students background knowledge, as well as a summary of the current status of high energy physics research
- Through exercises based on experiments give a realistic experience of how researchers work
- Create a meeting place with professors, young researchers and university students
- Offer a possibility to meet other Masterclass participants in Europe to compare results and discuss with experts at CERN
- Give a realistic picture of how research is carried out in international collaborations
- Present the local particle physics community and if possible, show local experiments

In Oslo the Masterclass became from the very beginning a group project. We deliberately involved master students both in lectures and as tutors during the exercises.

The majority of the participants came with their whole class, however there were also smaller groups of highly interested students. As interest has grown enormously, especially after the LHC start-up, consideration has been given to limit the target group to students who chose physics or can use the Masterclass in an assigned project.

#### Agenda

The international organizers provided a common agenda which started with 2-3 lectures during the morning session. Then there was a lunch where the organizers were present and available for questions and discussions. The afternoon was dedicated to an “experimental” part. The last hour was spent in a video conference connecting the 4-6 other Masterclass groups all over Europe and a moderator with his or her expert team at CERN.

During one of the years we also managed to organize, in collaboration with the Oslo Cyclotron Laboratory, a visit to their facility at the Department of Physics.

#### Lectures

There were usually three lectures covering:

- General introduction to particle physics. We found it also important to show the link to astrophysics and the evolution of the universe.
- Presentation of the experiments at CERN, mainly the LHC accelerator and the ATLAS experiment where the group is involved.
- Introduction to the experimental part covering the theory of  $Z^0$  decays where the anticipated results were presented as a hypothesis which was to be tested in the experiment. Students were also given a more concrete description of the DELPHI detector used in the “experiments” and examples of how to identify the various event and particle types.



### Experiment exercise

It was stressed that the exercise was actually a real experiment where we could not offer a ready made precise answer. Answers or clearer picture would only occur after careful observation of a large sample of collisions, discussions and combination of the measurements. The apparent lack of answer “at the back of the book” was for some participants a new and sometimes even frustrating experience.

The exercise was based on the educational project Hands-on-CERN [187] developed at the Stockholm House of Science [188]. It offered a well organized concise introduction to particle physics and collider experiments, as well as an event display of collisions in the DELPHI detector. The relevant pages were translated to Norwegian and a nearly complete local version of the project page was established [189].

The participants were working in groups of two or three. Over the years we experienced that a good introduction to the experiment session showing several examples was absolutely crucial. Many students found it also very useful to have “cheat sheet” which gave some key words for each event type and a schematic reminder of the various layers in a detector indicating which particles they register.

Each group of students had to analyze a batch of 100 collisions from the Hands-on-CERN collision library. As soon as they realized that most of the events were jet events they could identify the collisions quite quickly. However, each event collection contained some more ambiguous events. In such cases students experienced the importance of discussion and collaborative work.

The task of the students was to count the number of events in the four categories corresponding to  $Z^0$  decays:  $Z^0 \rightarrow e^+e^-$ ,  $Z^0 \rightarrow \mu^+\mu^-$ ,  $Z^0 \rightarrow \tau^+\tau^-$  and  $Z^0 \rightarrow q\bar{q}$ . Each year we experienced that the  $\tau$ -events were the most difficult. Many less typical jet events, often with relatively few tracks, were easily identified as  $\tau$ -events making these highly overrepresented in the final results.

The high number of participants required that each batch of 100 collisions was analyzed by more than one group. An average was therefore calculated for each of the 10 collision samples so the results which were to be presented during the video conference could fit the schema. Each of the participating Masterclass groups reported the numbers for a total of 1000 events.

Groups which finished early were encouraged to go through the events once more registering the number of two- and three-jet events. Based on the ratio of the two categories of events they could give an estimate of the strong coupling constant  $\alpha_S$ .

### Video linkup

The success of the video conference depended highly on the technical setup. Although it involved only simple technology like a normal web camera and standard audio devices, it was very sensitive to network problems. Several times it was experienced that one participating group could spoil the conference with bad audio setup or poor noise and echo suppression. We used a setup with two projectors, one for the main communication window showing the other participants, the second dedicated to a shared desktop broadcasted by the CERN team, which showed the result table and questions during the quiz.

The conference was divided in two parts. First results were gathered from all participating Masterclasses and the overall result was calculated followed by comments and discussion with the experts. The second part was a quiz where the answers were given

immediately after so that one could identify the winners of the day.

In the beginning we were excited about showing the high school students how communication technology is used to build international collaborations. However over the last year we have experienced fast development of social internet media and communities. Many students take these tools now for granted and find this way of communicating very familiar. We found that we no longer impress students with technology, however we were still able to challenge them by giving a realistic, good and meaningful experience of international collaborative research work.

### Feedback

The general feedback, both from students and teachers, has over the years been very positive. An evaluation questionnaire <sup>1</sup> showed that experiments in the physics classes are not very frequent and that computers are almost never used for data analysis. So although almost all students reported that they spent more than one hour per day in front of a computer, using it as a data analysis tool was new.

It was a surprise to see that many students said that the lectures were the part they liked most (73% in 2005, 53% in 2008). The exercises were second in popularity, while there were only very few who held the video linkup as their favourite.

The majority of participants answered that they at least partly agreed with the statements that the lectures were easy to understand and interesting. Most students found the level of Masterclass either as “exactly right” or a bit difficult. More than 50% answered that after the Masterclass they are more interested in physics in general.

It was interesting to note that most students at least partly supported the statement that they from attending the Masterclass had learned how scientific research was organized and carried out.

Also the feedback from the teachers has been very positive. They considered the project as relevant not only for the students, but also as an interesting update of their own knowledge.

Table 13.1 shows a summary of the answers about how the students liked the Masterclass.

### Outlook

After the LHC start-up in September 2008 we have experienced a great increase of interest in high energy physics and CERN. The book “Angels and Demons” by Dan Brown and later the film based on it has also been an unexpected but, still mostly positive factor. For the moment, the coordinated international Masterclass is continuing (the 2009 and 2010 events in Oslo have been organized by a new PhD student). Given the good feedback from the participants and the positive impact the effort has had in the group, there is a strong motivation to continue organizing such events in Oslo.

## 13.2 Department booth at National Science Fair 2008

### “What are you made of?”

The LHC start-up in September 2008 coincided with the yearly National Science Fair and the Department of Physics (UiO) decided to make LHC and ATLAS the theme of their

---

<sup>1</sup>The presented observations are based on the evaluation of the 2005 and 2008 event.

|      | How did you like the Masterclass you attended today? |     |     |     |                 |
|------|--|-----|-----|-----|-----------------|
| Year | very much<br>1                                       | 2   | 3   | 4   | not at all<br>5 |
| 2005 | 23%  | 46% | 23% | 9%  |                 |
| 2008 | 17%  | 33% | 38% | 13% |                 |

Table 13.1: General evaluation of the Masterclass in 2005 and 2008.



Figure 13.1: Cloud chamber [190] used at several exhibitions and outreach booths. The cooling dry ice is placed inside the black frame under the metal chamber floor. The cloth-covered brim of the chamber plexiglass box is moistured with alcohol and the alcohol vapor sinks to the bottom as it is being cooled. Incident particles leave “condense trails” very much like the familiar jet trails in the sky.

department booth. The author of the thesis was assigned by the Experimental Particle Physics group to be the coordinator of the project in collaboration with the department outreach officer and outreach committee.

Preparing a booth presenting particle physics is challenging as it is difficult to make it concrete and “hands-on”. The group did however develop some exhibits which worked very well.

In Oslo we have a detailed 1:40 ATLAS model which has been very successful. Presenting only pictures may often be confusing as many people have difficulties with imagining the experimental setup.

“Showing” particles is notoriously difficult. However, the department owns a simple cloud chamber operated with dry ice as cooling and alcohol vapour as active medium. It has been used with great success on many occasions. The working principle is easy to explain and people can follow the preparations of it. Cosmic rays and natural radiation is also often easier to grasp than the more exotic particles in collision experiments. The immediate experience of really “seeing” particle tracks in the cloud chamber has fascinated visitors, from children to retired researchers. Figure 13.1 shows a picture of the cloud chamber which was used at several outreach events. During the Science Fair we used a small round version with a diameter of approximately 15 cm.

Building on experience from several events, one of the most popular things is to offer some sort of activity, something one can do. If possible, it is even better if one in addition can get a souvenir. Trying to meet these requirements, we developed an activity that aimed at making the link between ourselves, something close and familiar, and the more

abstract world of particles. The idea was to start by weighting a person on a normal bathroom weight scale. The weight was fed into a script which divided the weight in different component like water, fat, proteins. In the next step the weight was given as number of atoms of various chemical elements. The journey continued to ever smaller particles giving the number of protons, neutrons and electrons and finally also number of fundamental particles, electrons, up- and down-quarks. This break-down was printed out with additional information about the experiments at CERN. Although very simple, the activity was a great success, especially since many people are concerned about their weight. This focus allowed us also to introduce the important question of mass in nature and the search for the Higgs particle as one of the main goals of the LHC experiments. As the title if the booth we therefor chose: “What are you made of?”.

Figure 13.2 shows an example of a “weight certificate”. The idea was worked out by the booth organizing committee while the programming was carried out by Eirik Gramstad.

### 13.3 Radio lecture: “There has to be something more”

Following up the enormous interest in CERN after the LHC start-up, the Norwegian National Broadcasting company invited the author of the thesis to hold a 26 minute long popular science lecture. Section 1.1, the general introduction, is partly based on material collected for the lecture. The introduction follows similar ideas, while the actual content is different. The main difference is that the introduction follows a historical route introducing the various particles and interactions.

One of the ideas of the radio lecture was not only to describe the current picture of the fundamental building blocks and interactions, but also give a flavor of the dynamics of the scientific work. The history of particle physics is full of fascinating examples of interaction between theory and experiment. Standing now on the threshold of a new era in high energy physics, it is challenging to explain to the general public how these enormous and expensive experiments are motivated and why there are such great expectations related to them. It can therefore be fruitful to present the SM in a historic perspective. On one hand we have its undeniable successes, on the other the clear evidence that it cannot be the final theory and full picture. This approach focuses on the motivation behind the LHC experiments and this is also the idea behind the title of the lecture: “There has to be something more”.

### 13.4 Outlook

In the Norwegian society and media there has over the past years been a relatively low interest in natural science, although indications are that this might be changing. This is somehow a paradox knowing the extensive impact technology and science have on our lives. Also the number of high school and university students choosing physics and mathematics has been alarmingly modest <sup>2</sup>. These subjects are often considered as “nerdy”, attaching to people in the field a rather suspicious social label.

It is tempting to ask if the focus of natural science teaching and outreach have been too focused on technology and applications, while often missing the deeper influence on our culture, social and political life, and the understanding of our world. I believe it is important to present natural science as part of our modern culture, not only through

---

<sup>2</sup>The latest numbers show that the trend might be changing. One of the reasons could be that the financial crisis has raised the interest for safe jobs.

## PARTIKKEL-ID

NAVN: Ola Nordmann

KLASSE: 3a

DETTE ER DU LAGET AV:

VISTE DU AT ...

29.25kg vann  
5.4kg fett  
9kg proteiner  
0.495kg RNA/DNA  
0.855kg annet

ELLER

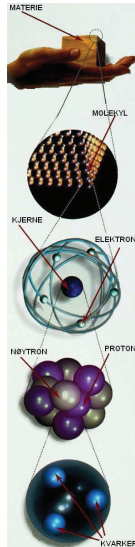
$2.68 \cdot 10^{27}$  hydrogenatomer  
 $1.03 \cdot 10^{27}$  oksygenatomer  
 $5.19 \cdot 10^{26}$  karbonatomer  
 $5.19 \cdot 10^{26}$  nitrogenatomer  
 $9.46 \cdot 10^{24}$  kalsiumatomer

ELLER

$1.46 \cdot 10^{28}$  protoner  
 $1.19 \cdot 10^{28}$  nøytroner  
 $1.46 \cdot 10^{28}$  elektroner

ELLER

$1.46 \cdot 10^{28}$  elektroner  
 $4.11 \cdot 10^{28}$  opp-kvarker  
 $3.84 \cdot 10^{28}$  ned-kvarker



... det jobber ca. 3000 mennesker på CERN fra 80 forskjellige land?

... på CERN beveger protonene seg i en ring med omkrets på 27 km, 100 meter under bakken?



... protonene er samlet i klumper med tusen milliarder protoner i hver klump?

... protonene beveger seg med en hastighet på 0.999999991 ganger lysets hastighet? Dvs. at de bruker 0.000100 s på en runde.

... at to protonklumper kolliderer hvert



0.45 kg (1%) av massen din kommer fra Higgs-mekanismen

44.55 kg (99%) av massen din kommer fra bevegelsen til kvarkene samt bindingene mellom partiklene

## STORE TALL

$10^6 \rightarrow 1.000.000$  (1 million)  
 $10^9 \rightarrow 1.000.000.000$  (1 milliard)  
 $10^{12} \rightarrow 1.000.000.000.000$  (1 billion)  
...  
 $10^{27} \rightarrow 1.000.000.000.000.000.000.000.000$   
(1000 kvadrillioner)



Gruppen for Eksperimentell Partikkelfysikk (EPF)  
Fysisk institutt  
Universitetet i Oslo  
<http://www.fys.uio.no>

CERN (European Organization for Nuclear Research)  
Sveits/Frankrike  
<http://www.cern.ch>



Figure 13.2: Mass break down calculated and handed out as a “weight certificate” at the Science Fair booth in September 2008. The idea was worked out in the organizing committee while the programing was carried out by Eirik Gramstad.

its applications, but also through its more philosophical impact. This view is based on the observation that many people find particle physics and CERN fascinating not only because of the spectacular experiments, but also because of the fundamental questions being asked. The relation between particle physics and our universe and its evolution is therefore an important part of the presentation of the field. The current mind-stretching concepts and challenging open questions remind us about the importance and value of fundamental research.

# Appendix A

## ATLAS computing

### A.1 Environment variables for ATLAS building

For the building procedure, the following variables must be set:

- **ATLAS\_ROOT**: Installation directory of the ATLAS-software itself. Many external packages can, for convenience, be installed here as well. While building the recommendation is to set this to  
*< installation directory just created > /dist/ < ATLASversion >.*
- **G4INSTALL**: Installation directory of Geant4. While building the recommendation is to set this to  
*< installation directory just created > /geant4/ < Geant4version >.*  
**G4SYSTEM** refers to the Linux environment, for example **Linux-gcc3**, and should not be changed.
- **ROOTSYS**: Installation directory of root. While building the recommendation is to set this to  
*< installation directory just created > /root/ < rootversion >.*
- **CERN\_LEVEL**: Version of the cern-libraries.
- **CERN**: Installation directory of the cern libraries. While building the recommendation is to set this to  
*< installation directory just created > /cern/\$CERN\_LEVEL.*
- **CMTROOT**: The recommendation is to set this to  
*< installation directory just created > /atlas/CMT/ < CMTversion >.*
- **CMTBIN**: Directory containing the binaries in the CMT-installation. If CMT is not patched in order to add the build platform, it must be set to **CMTBIN=Linux-i686a**.
- **SCRAM\_HOME**: Installation home of SCRAM. Can be set to **ATLAS\_ROOT**
- **JAVA\_HOME**: Home of a java installation on the build system.

Having defined these, one should set **PATH** and **LD\_LIBRARY\_PATH** to:

```
> PATH=$CERN/$CERN_LEVEL/bin:$JAVA_HOME/bin:$ATLAS_ROOT/bin
:$ROOTSYS/bin:$SCRAM_HOME/bin:$CMTROOT/$CMTBIN:$PATH

> LD_LIBRARY_PATH=$GCC_DIR/lib:$ATLAS_ROOT/lib:$ROOTSYS/lib
:$G4INSTALL/lib/$G4SYSTEM
```

## A.2 External packages

The ATLAS software consists of a large number of packages. Some of these are standard tools which are often part of the system and can be used directly if the versions required by ATLAS is compatible. Other are more HEP-specific like Monte Carlo generators. At CERN such HEP packages are maintained by the LHC Computing Grid (LCG) project and are found in the directory `/afs/cern.ch/sw/lcg` and in the ATLAS area `/afs/cern.ch/atlas/offline/external/`.

In addition, two software management packages are used: SCRAM (Software Configuration And Management) [191] is used by some of the LCG packages while CMT (Configuration Manager Tool) [119] is used by some other packages and by the ATLAS software. The software management packages CMT and SCRAM must be built and installed before a number of LCG and ATLAS packages can be built (see the dependencies section in Appendix A.3).

A part of the external software needed by ATLAS is managed and interfaced by the **External** package. Another part which is related to HEP-software is interfaced by the external package **LCGCMT**. In order to collect the correct versions of the source code for a release one has to browse these packages. There is a whole complex chain of dependencies. It starts from finding the release of the Gaudi [192] package which is a LHCb software used by Athena. This corresponds to a specific **LCGCMT** version which again has a package called **LCG\_Interface** which specifies the required versions of the software found in the **LCGCMT** area. The **LCG\_Interface** contains the CMT configuration files, the **requirements**-files which establish the link between the ATLAS software and the external packages.

A list of external packages and source code references is listed in Table A.1. All external packages were built and installed using the RPM tool kit. The source code was placed in the **SOURCES** directory while the RPM spec file was equipped with the required build commands. A CERN like software environment with the **ATLAS\_ROOT** as starting point was built by implementing link structure satisfying the expected paths defined in the CMT **requirements** files.

As there are many dependencies, the packages needed to be installed in a certain order. The full list of build and installation order can be found in next section.

## A.3 Installation of ATLAS external packages

The ATLAS software depends on a number of external packages which need to be installed in a certain order as shown in the list below.

- |              |               |
|--------------|---------------|
| 1. graphviz  | 5. pcre       |
| 2. cernlib   | 6. ulxmlrpcpp |
| 3. cascase   | 7. gsl        |
| 4. boost-jam | 8. AIDA       |



- |   |   |
|---|---|
| <ul style="list-style-type: none"> <li>9. root: depends on cernlib</li> <li>10. clhep</li> <li>11. CTVFMT</li> <li>12. phojet</li> <li>13. tauola</li> <li>14. scram</li> <li>15. cppunit</li> <li>16. cmake</li> <li>17. geant4</li> <li>18. xerces-c</li> <li>19. HTL</li> <li>20. castor</li> <li>21. HepUtilities: depends on AIDA</li> <li>22. boost: depends on boost-jam</li> <li>23. EvtGen: depends on CLHEP</li> <li>24. ICondDB</li> <li>25. TDAQ: depends on cmt and castor</li> <li>26. gccxml: depends on cmake</li> <li>27. Sherpa: depends on CLHEP</li> <li>28. CondDBMySQL: depends on mysql</li> </ul> | <ul style="list-style-type: none"> <li>29. mysql++: depends on mysql</li> <li>30. apache-ant</li> <li>31. SEAL: depends on scram, e2fsprogs, gccxml, boost, root, clhep, gsl, zlib, bz2lib, pcre, graphviz, and cppunit</li> <li>32. GENSER: depends on scram, cernlib, castor</li> <li>33. PI: depends on SEAL, scram, gccxml, boost, mysql++, mysql, python, root, xerces-c and pcre</li> <li>34. POOL: depends on SEAL, scram, gccxml, boost, mysql++, mysql, python, root, xerces-c and pcre</li> <li>35. COOL: depends on SEAL, POOL, boost, cppunit, gccxml, MyODBC, mysql, Oracle, sqlight, unixODBC, uuid, xerces-c, python, pcre</li> <li>36. Gaudi: depends on PI, SEAL, xerces-c, root, CLHEP, boost, python, HTL and cernlib</li> </ul> |
|---|---|

Table A.1 lists the versions and sources of the external software required by ATLAS 11.0.42, the last release which was distributed by the Oslo group using RPM.

## A.4 KitValidation

An example of a successful KitValidation output:

```
#####
##      Atlas Distribution Kit Validation Suite      ##
##      29-01-2006   v1.9.5-1                      ##
##                                                    ##
## Alessandro De Salvo  ##
#####
Testing AtlasRelease 11.0.4
athena executable                [PASSED]
athena shared libs               [PASSED]
Release shared libraries         [PASSED]
Release Simple Checks           [ OK ]
Athena Hello World              [ OK ]
...
Reconstruction Example (RecExCommon) [ OK ]
RecExToESD                      [ OK ]
DC3 Z -> e e Event Generation     [ OK ]
DC3 McAtNlo Event Generation     [ OK ]
DC3 Z -> e e G4 Simulation        [ OK ]
DC3 Z -> e e Digitization         [ OK ]
DC3 Z -> e e Reconstruction       [ OK ]
#####
##      Kit 11.0.4 Validation [ OK ]
#####
```

| Package                      | Version         | Source  |
|------------------------------|-----------------|---|
| atlasvirtual                 | 1.32-5          |   |
| perl-DBI                     | 2.81-88.7       |   |
| perl-CGI                     |                 |   |
| python2.3                    | 2.3.4           | rpm download site   |
| python2.3-devel              | 1.6.2-1         | <a href="http://archive.apache.org/dist/ant/source/">http://archive.apache.org/dist/ant/source/</a> |
| apache-ant                   | 2.3.0-4         |   |
| xerces-c-2.3.0-4             |                 |   |
| MySQL-shared                 | 4.0.18          | MySQL site  |
| MySQL-devel                  |                 |   |
| sqlite                       |                 |   |
| sqlite-devel                 |                 |   |
| oracle-instantclient-basic   |                 |   |
| oracle-instantclient-devel   |                 |   |
| oracle-instantclient-sqlplus | 10.1.0.3        | Oracle site   |
| cppunit                      |                 |   |
| cmake                        |                 |   |
| gccxml                       | 0.6.0           |   |
| scram                        | V0.20.0         | from CVS anonymous@spitools.cvs.cern.ch   |
| CMT                          | v1r16p20040901  | From the CMT site   |
| ulxmlrpcpp                   | 1.3.0           |   |
| unixODBC                     |                 |   |
| unixODBC-devel               | 2.2.10          |   |
| MyODBC                       | 3.51.10         |   |
| AVTOOLS                      | 2.2-1           | afs   |
| ConddbMySQL                  | 0.5.6           |   |
| boost-jam                    | 3.1.7           | <a href="http://sourceforge.net/">http://sourceforge.net/</a>                                       |
| boost                        | 1.31.0          | <a href="http://sourceforge.net/">http://sourceforge.net/</a>                                       |
| cernlib                      | 2003            | <a href="http://sourceforge.net/">http://sourceforge.net/</a>                                       |
| castor                       | 1.7.1.5-1       | <a href="http://sourceforge.net/">http://sourceforge.net/</a>                                       |
| CLHEP18                      | 1.8.2.1-2.atlas | <a href="http://sourceforge.net/">http://sourceforge.net/</a>                                       |

|              |                     |   |
|--------------|---------------------|---|
| CLHEP19      | 1.9.1.2-spi1.1      | wwasd.web.cern.ch/wwasd/lhc++/clhep/INSTALLATION/clhep-1.9.html           |
| geant4       | 4-7.0.p01-1.8.2.1.1 | geant4.cern.ch/geant4/source/source/                                      |
| CTVMFT       | v1.0-1              | /afs/cern.ch/atlas/offline/external/CTVMFT                                |
| toprex       | TRX406-1            |   |
| tauola-cleo  |                     | /afs/cern.ch/atlas/offline/external/tauola/tauola-cleo/                   |
| bases        | 5-1                 | ftp://ftp.kek.jp/kek/minami/bases/basesv5.1.tar.gz                        |
| cascade      | 1.2001b-1           |   |
| phojet       | 112135              | /afs/cern.ch/atlas/offline/external/phojet/                               |
| GENSER       | 0.3.0               | /afs/cern.ch/sw/lcg/app/releases/GENSER/                                  |
| sherpa       | V1.0.3              | /afs/cern.ch/atlas/offline/external/sherpa/                               |
| herwig6      | 505c-1              |   |
| HTL          | 1.3.2.1-13          |   |
| graphviz     | 2.0-1               |   |
| g95          | 3.5.0-1             |   |
| axis         | 1.1-1               |   |
| EvtGen-alpha | 00.10.22-2          | /afs/cern.ch/atlas/offline/external/EvtGen/                               |
| tdaq         | 01.01.00            | http://atddoc.cern.ch/cmt/releases/                                       |
| gsl          | 1.5-1               |   |
| pcr          |                     |   |
| pre-devel    | 4.4-1               |   |
| root         | 4.03.02             | ftp://root.cern.ch/root/root.v.source.tar.gz                              |
| SEAL         | 1.6.1               | /afs/cern.ch/sw/lcg/app/seal/ & http://lcgapp.cern.ch/project/            |
| mysql++      | 1.7.9-7             |   |
| AIDA         | 3.2.1               | ftp://ftp.slac.stanford.edu/software/freehep/AIDA                         |
| HepUtilities | 0.7.1.1-2           | /afs/cern.ch/sw/lhcxx/share/HepUtilities/0.7.1.1/                         |
| POOL         | 2.0.1               | /afs/cern.ch/sw/lcg/app/pool/ & http://lcgapp.cern.ch/project/            |
| PI           | 1.3.0               | http://pi.cvs.cern.ch/cgi-bin/pi.cgi/pi/ & http://lcgapp.cern.ch/project/ |
| Gaudi        | 0.14.6.14.pool201   | /afs/cern.ch/atlas/offline/external/Gaudi                                 |

Table A.1: ATLAS external software.

| Field        | Content   |
|--------------|---|
| TASKID       | Unique task identification                                  |
| TASKNAME     | Human readable name which refers to certain physics process |
| STATUS       | Task status   |
| CPUCOUNT     | Expected CPU usage (kSI2Ks)                                 |
| DISKCOUNT    | Expected disk usage (MB)                                    |
| RAMCOUNT     | Expected memory usage (MB)                                  |
| TRANSHOME    | The production cache to be used AtlasProduction_14.2.10.1   |
| TRANSPATH    | The job transformed for be used                             |
| GRID         | The grid to which the task is assigned                      |
| TIER         | The Tier (cloud) to which the task is assigned to           |
| TASKPRIORITY | Priority of the task 100 - 999                              |

Table A.2: Proddb table ETASK

## A.5 ATLAS production database Proddb

The ATLAS Production Database (Proddb) is a central component of the Production System which is described in Section 5.7.2.

The Proddb job definition and bookkeeping are organized in a schema consisting of 3 tables: ETASK, EJOBDEF and EJOBEXE. These will now be shortly described. The fields selected for a more detailed description are relevant for the implementation of the job handling in the executor as well as for the management of the production on the NorduGrid resources.

The global task specification is kept in the ETASK table. Table A.2 summarizes the most important fields.

The EJOBDEF, Table A.3, contains all information necessary to define the particular jobs within a task. On the job level one needs to specify parameters like input, output and log files of the job, as well as which events are to be processed. The JOBPARS field contains arguments for the job transform script, like for example number of events to be processed, random seed and detector geometry. The list of arguments vary with the task type.

Table A.4, the EJOBEXE, contains one entry per job which is kept up-to-date following as many attempts as the job might go through, until it ends up in one of the final states (like DONE or ABORTED). The EJOBDEF record keeps only the current status of the job, while its “history”, meaning the information about the subsequent attempts is kept in the EJOBEXE table. The EJOBEXE table contains, in addition to some of the same fields as EJOBDEF table, a more detailed information about the particular attempt of the job. These include the executor which handled it, exactly where it was run, what was the actual resource usage (CPU/wall time). In case of failure, two sets of error codes and acronyms are filled in, one from the production system itself and one provided by the job transforms. There is also a field for a short error message. This information is crucial for debugging and monitoring of the computing resource performance. The content of the EJOBEXE table is of great interest for people running the production. A good information and monitoring system which allows to spot problems early is the key to high efficiency.

| Field         | Content   |
|---------------|---|
| CODIFIED      | Unique job identification   |
| JOBNAME       | Human readable name, same as the task name, plus PARTNR                     |
| TASKFK        | Unique task identification, TASKFK == TASKID in ETASK table                 |
| CURRENTSTATUS | Status  |
| LASTATTEMPT   | Number indicating how many attempts the job has had                         |
| MAXATTEMPT    | Maximun number of attempts  |
| PRIORITY      | Priority of the job (xml in EJOBDEFBIG)                                     |
| JOBOUTPUTS    | Output files (xml in EJOBDEFBIG)  |
| JOBLOGS       | Log files (xml in EJOBDEFBIG)   |
| JOBPARS       | Lists specific parameters of a job (e.g., input files, random numbers ... ) |
| LOCKEDBY      | Indicate the executor handling the job                                      |
| TRANSINFO     | Atlas software and job transform to be used                                 |
| SOURCETIER    | Tier where the input data are stored  |

Table A.3: Proddb table EJOBDEF

| Field           | Content  |
|-----------------|--|
| JOBEXEID        | Unique attempt identification                              |
| JOBDEFFK        | Unique job identification, JOBDEFFK == JOBDEFID in EJOBDEF |
| ATTEMPTNR       | Number for this particular execution                       |
| EXECUTOR        | The identifier of the executor which handled the job       |
| JOBSTATUS       | Staus  |
| JOBNATIVESTATUS | Status native to the executing system (grid)               |
| EXECLUSTER      | Cluster which ran the job                                  |
| CPUCOUNT        | Used CPU-time (s)  |
| ERRORCODE       | Error code according to a schema                           |
| ERRORACRONYM    | Tag indicating the nature of the failure                   |
| ERRORTXT        | Error message provided by the system                       |
| TRFCODE         | Job transform generated error code                         |
| TRFACRONYM      | Corresponding tag indicating the nature of the failure     |

Table A.4: Proddb table EJOBEXE. The table contains also fields like TASKFK, JOBNAME, PARTNR which are the same as in the EJOBDEF table



# Appendix B

## SUSY branching fractions

The tables below give an overview of the main decay channels for supersymmetric particles in the four studied mSUGRA benchmark points introduced in Sections 6.3.1–6.3.4. The listing shows only the main decay channels and in most cases the branching fractions do not sum up to 100%. Leptons, electrons and muons, are represented by  $l$  and the leptonic branching fraction is a sum for both flavors.

| Sparticle          | Decay                       | BR[%] | Sparticle          | Decay                  | BR[%] | Sparticle          | Decay                       | BR[%] |
|--------------------|-----------------------------|-------|--------------------|------------------------|-------|--------------------|-----------------------------|-------|
| $\tilde{\chi}_1^+$ | $\tilde{\nu}_l \nu_l$       | 42    | $\tilde{\chi}_2^+$ | $\tilde{\chi}_2^0 W^+$ | 28    | $\tilde{q}_L$      | $\tilde{\chi}_1^+ q'$       | 64    |
|                    | $\tilde{\nu}_\tau \tau$     | 26    |                    | $\tilde{\chi}_1^+ Z$   | 24    |                    | $\tilde{\chi}_2^0 q$        | 32    |
|                    | $\tilde{\tau} \nu_\tau$     | 18    |                    | $\tilde{\chi}_1^+ h$   | 20    |                    | $\tilde{\chi}_1^0 q$        | 99    |
|                    |                             |       |                    | $\tilde{l}_L \nu_l$    | 9     |                    | $\tilde{\chi}_1^0 l$        | 100   |
| $\tilde{\chi}_2^0$ | $\tilde{\nu}_l \nu_l$       | 37    | $\tilde{\chi}_3^0$ | $\tilde{\chi}_1^+ W^-$ | 58    | $\tilde{\ell}_R$   | $\tilde{\chi}_1^0 l$        | 100   |
|                    | $\tilde{\nu}_\tau \nu_\tau$ | 23    |                    | $\tilde{\chi}_2^0 Z$   | 24    | $\tilde{\tau}_1$   | $\tilde{\chi}_1^0 \tau$     | 100   |
|                    | $\tilde{\tau}^+ \tau^-$     | 21    |                    | $\tilde{\chi}_1^0 Z$   | 11    | $\tilde{\nu}_\tau$ | $\tilde{\chi}_1^0 \nu_\tau$ | 98    |
| $\tilde{\chi}_4^0$ | $\tilde{\chi}_1^+ W^-$      | 51    | $\tilde{g}$        | $\tilde{q} \bar{q}$    | 37    | $\tilde{\nu}_l$    | $\tilde{\tau}_1 W^-$        | 2     |
|                    | $\tilde{\chi}_2^0 h$        | 17    |                    | $\tilde{t}_1 \bar{t}$  | 18    |                    | $\tilde{\chi}_1^0 \nu_l$    | 100   |
|                    | $\tilde{\nu}_l \nu_l$       | 9     |                    | $\tilde{b}_1 \bar{b}$  | 15    |                    |                             |       |
| $\tilde{t}_1$      | $\tilde{\chi}_1^+ b$        | 49    | $\tilde{b}$        | $\tilde{\chi}_1^- t$   | 34    |                    |                             |       |
|                    | $\tilde{\chi}_1^0 t$        | 23    |                    | $\tilde{\chi}_2^- t$   | 25    |                    |                             |       |
|                    | $\tilde{\chi}_2^0 t$        | 15    |                    | $\tilde{\chi}_2^0 b$   | 22    |                    |                             |       |

Table B.1: Branching ratios of particles in the SU1 benchmark point. The listing for some of the particles omits the lowest branching fractions.  $l$  stands for the lightest leptons  $e$  and  $\mu$ . SU1 is described in Section 6.3.1.

## B. SUSY branching fractions

| Sparticle                             | Decay                                      | BR[%] | Sparticle          | Decay                  | BR[%] |
|---------------------------------------|--|-------|--------------------|------------------------|-------|
| $\tilde{\chi}_1^-$                    | $\tilde{\chi}_1^0 q \bar{q}'$              | 67    | $\tilde{\chi}_2^+$ | $\tilde{\chi}_2^0 W$   | 37    |
|                                       | $\tilde{\chi}_1^0 \nu_l l$                 | 22    |                    | $\tilde{\chi}_1^+ Z$   | 33    |
|                                       | $\tilde{\chi}_1^0 \nu_\tau \tau$           | 11    |                    | $\tilde{\chi}_3^0 W$   | 15    |
| $\tilde{\chi}_2^0 / \tilde{\chi}_3^0$ | $\tilde{\chi}_1^0 q \bar{q}$               | 51    | $\tilde{g}$        | $\tilde{\chi}_1^+ t b$ | 26    |
|                                       | $\tilde{\chi}_1^0 b \bar{b}$               | 13    |                    | $\tilde{\chi}_2^+ t b$ | 19    |
|                                       | $\tilde{\chi}_1^0 \nu_l \bar{\nu}_l$       | 13    |                    | $\tilde{\chi}_3^0 t t$ | 10    |
|                                       | $\tilde{\chi}_1^0 \nu_\tau \bar{\nu}_\tau$ | 6     |                    | $\tilde{\chi}_2^0 t t$ | 9     |
|                                       | $\tilde{\chi}_1^0 l^+ l^-$                 | 5     |                    | $\tilde{\chi}_2^+ u d$ | 7     |
|                                       |  |       |                    | $\tilde{\chi}_4^0 t t$ | 5     |
| $\tilde{\chi}_4^0$                    | $\tilde{\chi}_1^+ W$                       | 84    |                    | $\tilde{\chi}_1^+ u d$ | 4     |
|                                       | $\tilde{\chi}_3^0 Z$                       | 8     |                    |                        |       |
|                                       | $\tilde{\chi}_2^0 h$                       | 7     |                    |                        |       |

Table B.2: Branching ratios of sparticles in the SU2 benchmark point. SU2 is described in Section 6.3.2.

| Sparticle          | Decay                       | BR[%] | Sparticle          | Decay                  | BR[%] | Sparticle        | Decay                       | BR[%] |
|--------------------|-----------------------------|-------|--------------------|------------------------|-------|------------------|-----------------------------|-------|
| $\tilde{\chi}_1^+$ | $\tilde{\tau}_1 \nu_\tau$   | 68    | $\tilde{\chi}_2^+$ | $\tilde{\chi}_2^0 W^+$ | 28    | $\tilde{q}_L$    | $\tilde{\chi}_1^+ q$        | 66    |
|                    | $\tilde{\chi}_1^0 W$        | 29    |                    | $\tilde{\chi}_1^+ Z$   | 25    |                  | $\tilde{\chi}_2^0 q$        | 33    |
|                    | $\tilde{\nu}_l l$           | 2     |                    | $\tilde{\chi}_1^+ h$   | 20    | $\tilde{q}_R$    | $\tilde{\chi}_1^0 q$        | 99    |
| $\tilde{\chi}_2^0$ | $\tilde{\tau}_1 \tau$       | 75    | $\tilde{\chi}_3^0$ | $\tilde{\chi}_1^+ W$   | 61    | $\tilde{\ell}_L$ | $\tilde{\chi}_1^0 l$        | 82    |
|                    | $\tilde{l}_R l$             | 18    |                    | $\tilde{\chi}_2^0 Z$   | 25    |                  | $\tilde{\chi}_1^+ \nu_l$    | 11    |
|                    | $\tilde{\chi}_1^0 Z$        | 3     |                    | $\tilde{\chi}_1^0 Z$   | 10    |                  | $\tilde{\chi}_2^0 l$        | 6     |
| $\tilde{\chi}_4^0$ | $\tilde{\chi}_1^+ W$        | 56    | $\tilde{g}$        | $\tilde{q}_R q$        | 36    | $\tilde{\ell}_R$ | $\tilde{\chi}_1^0 l$        | 100   |
|                    | $\tilde{\chi}_2^0 h$        | 18    |                    | $\tilde{q}_L q$        | 23    | $\tilde{\tau}_1$ | $\tilde{\chi}_1^0 \tau$     | 100   |
|                    | $\tilde{\chi}_1^0 h$        | 7     |                    | $\tilde{t}_1 t$        | 16    | $\tilde{\tau}_2$ | $\tilde{\chi}_1^0 \tau$     | 81    |
|                    | $\tilde{\nu}_l \nu_l$       | 6     |                    | $\tilde{b}_1 b$        | 15    |                  | $\tilde{\chi}_1^+ \nu_\tau$ | 12    |
|                    | $\tilde{\nu}_\tau \nu_\tau$ | 3     |                    | $\tilde{b}_2 b$        | 9     |                  | $\tilde{\chi}_2^0 \tau$     | 6     |
|                    |                             |       |                    |                        |       |                  |                             |       |
| $\tilde{t}_1$      | $\tilde{\chi}_1^+ b$        | 65    | $\tilde{b}_1$      | $\chi_1^- t$           | 36    |                  |                             |       |
|                    | $\tilde{\chi}_1^0 t$        | 25    |                    | $\tilde{t}_1 W^-$      | 35    |                  |                             |       |
|                    | $\tilde{\chi}_2^0 t$        | 10    |                    | $\tilde{\chi}_2^0 b$   | 26    |                  |                             |       |

Table B.3: Branching ratios of sparticles in the SU3 benchmark point. SU3 is described in Section 6.3.3.

| Sparticle          | Decay                              | BR[%] | Sparticle          | Decay                         | BR[%] | Sparticle                        | Decay                       | BR[%] |
|--------------------|------------------------------------|-------|--------------------|-------------------------------|-------|----------------------------------|-----------------------------|-------|
| $\tilde{\chi}_1^+$ | $\tilde{\chi}_1^0 q \bar{q}'$      | 65    | $\tilde{\chi}_2^+$ | $\tilde{t}_1 b$               | 46    | $\tilde{q}_L$                    | $\tilde{\chi}_1^+ q'$       | 66    |
|                    | $\tilde{\chi}_1^0 \nu_l l^+$       | 23    |                    | $\tilde{\chi}_2^0 W^+$        | 19    |                                  | $\tilde{\chi}_2^0 q$        | 32    |
|                    | $\tilde{\chi}_1^0 \nu_\tau \tau^+$ | 12    |                    | $\tilde{\chi}_1^+ Z$          | 15    | $\tilde{q}_R$                    | $\tilde{\chi}_1^0 q$        | 98    |
| $\tilde{\chi}_2^0$ | $\tilde{\chi}_1^0 q \bar{q}$       | 56    |                    | $\tilde{\chi}_1^+ h$          | 11    | $\tilde{\ell}_L$                 | $\tilde{\chi}_1^- \nu_l$    | 55    |
|                    | $\tilde{\chi}_1^0 b \bar{b}$       | 18    | $\tilde{\chi}_3^0$ | $\tilde{\chi}_1^+ W^-$        | 62    |                                  | $\tilde{\chi}_2^0 l$        | 34    |
|                    | $\tilde{\chi}_1^0 \tau^+ \tau^-$   | 8     |                    | $\tilde{\chi}_2^0 Z$          | 20    |                                  | $\tilde{\chi}_1^0 l$        | 10    |
|                    |                                    |       |                    | $\tilde{\chi}_1^0 Z$          | 10    | $\tilde{\ell}_R$                 | $\tilde{\chi}_1^0 l$        | 98    |
| $\tilde{\chi}_4^0$ | $\tilde{\chi}_1^+ W$               | 59    | $\tilde{g}$        | $\tilde{t}_1 \bar{t}$         | 57    | $\tilde{\tau}_1$                 | $\tilde{\chi}_1^0 \tau$     | 72    |
|                    | $\tilde{\chi}_2^0 h$               | 15    |                    | $\tilde{b}_1 \bar{b}$         | 39    |                                  | $\tilde{\chi}_1^- \nu_\tau$ | 17    |
|                    | $\tilde{\chi}_1^0 h$               | 7     |                    | $\tilde{\chi}_1^+ q \bar{q}'$ | 2     |                                  | $\tilde{\chi}_2^0 \tau$     | 11    |
| $\tilde{t}_1$      | $\tilde{\chi}_1^+ b$               | 100   | $\tilde{b}$        | $\tilde{t}_1 W^-$             | 52    | $\tilde{\nu}_\tau / \tilde{\nu}$ | $\tilde{\chi}_1^+ l^-$      | 58    |
|                    |                                    |       |                    | $\tilde{\chi}_2^0 b$          | 25    |                                  | $\tilde{\chi}_2^0 \nu_l$    | 23    |
|                    |                                    |       |                    | $\tilde{\chi}_1^- t$          | 20    |                                  | $\tilde{\chi}_1^0 \nu_l$    | 19    |

Table B.4: Branching ratios of sparticles in the SU4 benchmark point. The listing for some of the sparticles omits the lowest branching fractions.  $l$  stands for the lightest leptons  $e$  and  $\mu$ . SU4 is described in Section 6.3.4.



# Appendix C

## List of 10 TeV MC samples

The tables below give a more detailed listing of signal and background MC samples used in Chapter 9 and 10. The SM and SUx samples correspond to those in Table 9.3. Table C.2 lists the  $p_T$  ranges associated with the JX tags used in the QCD di-jet sample names.

| Dataset                              | samp ID | # ev.   | AOD       | $\sigma_{LO}$ [pb] | $\sigma_{NLO}$ [pb] | Scale f. | $\int dt \mathcal{L}$ [fb <sup>-1</sup> ] |
|--------------------------------------|---------|---------|-----------|--------------------|---------------------|----------|---|
| SU1                                  | 105401  | 10000   |           | 2.42               | 3.22                | 0.064    | 3106.9                                    |
| SU3                                  | 105403  | 14999   |           | 5.46               | 7.26                | 0.097    | 2065.5                                    |
| SU4                                  | 106400  | 168802  |           | 107.40             | 147.14              | 0.174    | 1147.2                                    |
| Zgam                                 | 105120  | 23989   |           | 11.09              | 14.42               | 0.120    | 1663.9                                    |
| $t\bar{t}$                           | 105200  | 2266954 |           | 205.48             | 205.48              | 0.025    | 8140.4                                    |
| Wt                                   | 105500  | 19963   |           | 14.41              | 20.17               | 0.202    | 989.5                                     |
| Singl top. t-chan                    | 105502  | 29961   |           | 41.12              | 40.30               | 0.318    | 638.2                                     |
| WPhoton10                            | 105909  | 24991   |           | 22.42              | 29.15               | 0.233    | 1361.0                                    |
| $W^+W^- \rightarrow e\nu e\nu$       | 105921  | 24981   |           | 0.83               | 0.83                | 0.010    | 20832.2                                   |
| $W^+W^- \rightarrow e\nu \mu\nu$     | 105922  | 24958   |           | 0.83               | 0.83                | 0.009    | 20954.2                                   |
| $W^+W^- \rightarrow e\nu \tau\nu$    | 105923  | 24984   |           | 0.83               | 0.83                | 0.009    | 20773.1                                   |
| $W^+W^- \rightarrow \mu\nu \mu\nu$   | 105924  | 24973   |           | 0.83               | 0.83                | 0.009    | 20940.9                                   |
| $W^+W^- \rightarrow \mu\nu e\nu$     | 105925  | 25000   |           | 0.83               | 0.83                | 0.009    | 20831.0                                   |
| $W^+W^- \rightarrow \mu\nu \tau\nu$  | 105926  | 24976   |           | 0.83               | 0.83                | 0.009    | 21108.7                                   |
| $W^+W^- \rightarrow \tau\nu \tau\nu$ | 105927  | 24962   |           | 0.83               | 0.83                | 0.009    | 20797.2                                   |
| $W^+W^- \rightarrow \tau\nu e\nu$    | 105928  | 23922   |           | 0.83               | 0.83                | 0.009    | 20135.7                                   |
| $W^+W^- \rightarrow \tau\nu \mu\nu$  | 105929  | 24950   |           | 0.83               | 0.83                | 0.009    | 20860.0                                   |
| $ZZ \rightarrow llll$                | 105931  | 20000   |           | 0.04               | 0.04                | 0.001    | 329918.7                                  |
| $ZZ \rightarrow ll\nu\nu$            | 105932  | 19998   |           | 0.25               | 0.25                | 0.004    | 54915.9                                   |
| $W^+Z \rightarrow l\nu ll$           | 105941  | 25000   |           | 0.26               | 0.26                | 0.003    | 69013.7                                   |
| $W^+Z \rightarrow qqll$              | 105942  | 5000    |           | 0.83               | 0.83                | 0.047    | 4388.5                                    |
| $W^-Z \rightarrow l\nu ll$           | 105971  | 19957   |           | 0.16               | 0.16                | 0.002    | 93479.1                                   |
| $W^-Z \rightarrow qqll$              | 105972  | 5000    |           | 0.49               | 0.49                | 0.025    | 7486.2                                    |
| Wbb+Np0                              | 106280  | 15500   |           | 5.12               | 6.25                | 0.081    | 3308.6                                    |
| Wbb+Np1                              | 106281  | 15457   |           | 4.99               | 6.09                | 0.079    | 5770.5                                    |
| Wbb+Np2                              | 106282  | 8953    |           | 2.91               | 3.55                | 0.079    | 9006.5                                    |
| Wbb+Np3                              | 106283  | 5000    |           | 1.61               | 1.96                | 0.079    | 9428.0                                    |
| QcdbbJ2+Np0                          | 107300  | 137258  | 137665.26 | 167951.62          | 244.724             |          | 0.8                                       |
| QcdbbJ2+Np1                          | 107301  | 193184  | 193821.09 | 236461.73          | 244.805             |          | 0.8                                       |
| QcdbbJ2+Np2                          | 107302  | 53956   | 53806.52  | 65643.95           | 243.324             |          | 0.8                                       |
| QcdbbJ2+Np3                          | 107303  | 3500    | 13470.88  | 16434.47           | 939.113             |          | 0.2                                       |
| QcdbbJ3+Np0                          | 107305  | 5464    | 5398.06   | 6585.63            | 241.055             |          | 0.8                                       |
| QcdbbJ3+Np1                          | 107306  | 26920   | 27239.62  | 33232.34           | 246.897             |          | 0.8                                       |
| QcdbbJ3+Np2                          | 107307  | 18459   | 18591.99  | 22682.23           | 245.758             |          | 0.8                                       |
| QcdbbJ3+Np3                          | 107308  | 9472    | 9460.53   | 11541.85           | 243.705             |          | 0.8                                       |

| Dataset                         | samp ID | # ev.   | AOD      | $\sigma_{LO}$ [pb] | $\sigma_{+NLO}$ [pb] | Scale f. | $\int dt \mathcal{L}$ [fb <sup>-1</sup> ] |
|---------------------------------|---------|---------|----------|--------------------|----------------------|----------|---|
| QcdbbJ4+Np0                     | 107310  | 1464    | 147.90   | 180.44             | 24.650               | 8.1      |   |
| QcdbbJ4+Np1                     | 107311  | 11000   | 1078.60  | 1315.89            | 23.925               | 8.4      |   |
| QcdbbJ4+Np2                     | 107312  | 14500   | 1430.00  | 1744.60            | 24.063               | 8.3      |   |
| QcdbbJ4+Np3                     | 107313  | 10491   | 1021.20  | 1245.86            | 23.751               | 8.4      |   |
| QcdbbJ4+Np4                     | 107314  | 7000    | 706.50   | 861.93             | 24.627               | 8.1      |   |
| QcdbbJ5+Np0                     | 107315  | 1000    | 3.20     | 3.90               | 0.781                | 256.1    |   |
| QcdbbJ5+Np1                     | 107316  | 8000    | 25.20    | 30.74              | 0.769                | 260.2    |   |
| QcdbbJ5+Np2                     | 107317  | 15429   | 50.00    | 61.00              | 0.791                | 252.2    |   |
| QcdbbJ5+Np3                     | 107318  | 15967   | 52.90    | 64.54              | 0.808                | 247.4    |   |
| QcdbbJ5+Np4                     | 107319  | 16982   | 55.50    | 67.71              | 0.797                | 250.8    |   |
| Z $\rightarrow$ ee+Np0          | 107650  | 269280  | 898.20   | 1095.80            | 0.814                | 245.7    |   |
| Z $\rightarrow$ ee+Np0          | 107650  | 269280  | 898.20   | 1095.80            | 0.814                | 245.7    |   |
| Z $\rightarrow$ ee+Np1          | 107651  | 61767   | 206.60   | 252.05             | 0.816                | 245.1    |   |
| Z $\rightarrow$ ee+Np2          | 107652  | 216945  | 72.50    | 88.45              | 0.082                | 2452.7   |   |
| Z $\rightarrow$ ee+Np3          | 107653  | 63412   | 21.10    | 25.74              | 0.081                | 2463.4   |   |
| Z $\rightarrow$ ee+Np4          | 107654  | 18470   | 6.00     | 7.32               | 0.079                | 2523.2   |   |
| Z $\rightarrow$ ee+Np5          | 107655  | 5500    | 1.70     | 2.07               | 0.075                | 2651.9   |   |
| Z $\rightarrow$ $\mu\mu$ +Np0   | 107660  | 270098  | 900.20   | 1098.24            | 0.813                | 245.9    |   |
| Z $\rightarrow$ $\mu\mu$ +Np1   | 107661  | 61936   | 205.20   | 250.34             | 0.808                | 247.4    |   |
| Z $\rightarrow$ $\mu\mu$ +Np2   | 107662  | 207173  | 69.40    | 84.67              | 0.082                | 2446.9   |   |
| Z $\rightarrow$ $\mu\mu$ +Np3   | 107663  | 64956   | 21.60    | 26.35              | 0.081                | 2464.9   |   |
| Z $\rightarrow$ $\mu\mu$ +Np4   | 107664  | 18470   | 6.10     | 7.44               | 0.081                | 2481.9   |   |
| Z $\rightarrow$ $\mu\mu$ +Np5   | 107665  | 5471    | 1.70     | 2.07               | 0.076                | 2637.9   |   |
| Z $\rightarrow$ $\tau\tau$ +Np0 | 107670  | 270649  | 902.70   | 1101.29            | 0.814                | 245.8    |   |
| Z $\rightarrow$ $\tau\tau$ +Np1 | 107671  | 61928   | 209.30   | 255.35             | 0.825                | 242.5    |   |
| Z $\rightarrow$ $\tau\tau$ +Np2 | 107672  | 210234  | 70.20    | 85.64              | 0.081                | 2454.7   |   |
| Z $\rightarrow$ $\tau\tau$ +Np3 | 107673  | 63434   | 21.10    | 25.74              | 0.081                | 2464.2   |   |
| Z $\rightarrow$ $\tau\tau$ +Np4 | 107674  | 18500   | 6.00     | 7.32               | 0.079                | 2527.3   |   |
| Z $\rightarrow$ $\tau\tau$ +Np5 | 107675  | 5479    | 1.70     | 2.07               | 0.076                | 2641.8   |   |
| W $\rightarrow$ e $\nu$ +Np0    | 107680  | 1307302 | 10184.70 | 12425.33           | 1.901                | 105.2    |   |
| W $\rightarrow$ e $\nu$ +Np1    | 107681  | 260924  | 2112.30  | 2577.01            | 1.975                | 101.3    |   |
| W $\rightarrow$ e $\nu$ +Np2    | 107682  | 777754  | 676.00   | 824.72             | 0.212                | 943.1    |   |
| W $\rightarrow$ e $\nu$ +Np3    | 107683  | 224574  | 203.30   | 248.03             | 0.221                | 905.4    |   |
| W $\rightarrow$ e $\nu$ +Np4    | 107684  | 58872   | 56.10    | 68.44              | 0.233                | 860.2    |   |
| W $\rightarrow$ e $\nu$ +Np5    | 107685  | 16992   | 16.60    | 20.25              | 0.238                | 839.0    |   |
| W $\rightarrow$ $\mu\nu$ +Np0   | 107690  | 1328494 | 10125.70 | 12353.35           | 1.860                | 107.5    |   |
| W $\rightarrow$ $\mu\nu$ +Np1   | 107691  | 268747  | 2155.50  | 2629.71            | 1.957                | 102.2    |   |
| W $\rightarrow$ $\mu\nu$ +Np2   | 107692  | 780326  | 682.30   | 832.41             | 0.213                | 937.4    |   |
| W $\rightarrow$ $\mu\nu$ +Np3   | 107693  | 211344  | 202.00   | 246.44             | 0.233                | 857.6    |   |
| W $\rightarrow$ $\mu\nu$ +Np4   | 107694  | 57928   | 55.50    | 67.71              | 0.234                | 855.5    |   |
| W $\rightarrow$ $\mu\nu$ +Np5   | 107695  | 16975   | 16.30    | 19.89              | 0.234                | 853.6    |   |
| W $\rightarrow$ $\tau\nu$ +Np0  | 107700  | 1327107 | 10178.30 | 12417.53           | 1.871                | 106.9    |   |
| W $\rightarrow$ $\tau\nu$ +Np1  | 107701  | 263827  | 2106.90  | 2570.42            | 1.949                | 102.6    |   |
| W $\rightarrow$ $\tau\nu$ +Np2  | 107702  | 791658  | 672.80   | 820.82             | 0.207                | 964.5    |   |
| W $\rightarrow$ $\tau\nu$ +Np3  | 107703  | 222162  | 202.70   | 247.29             | 0.223                | 898.4    |   |
| W $\rightarrow$ $\tau\nu$ +Np4  | 107704  | 58765   | 55.30    | 67.47              | 0.230                | 871.0    |   |
| W $\rightarrow$ $\tau\nu$ +Np5  | 107705  | 15913   | 17.00    | 20.74              | 0.261                | 767.3    |   |

| Dataset                           | samp ID | # ev. AOD | $\sigma_{LO}$ [pb] | $\sigma_{+NLO}$ [pb] | Scale f. | $\int dt \mathcal{L}$ [fb $^{-1}$ ] |
|-----------------------------------|---------|-----------|--------------------|----------------------|----------|-------------------------------------|
| Z $\rightarrow$ eebb+Np0          | 109300  | 299757    | 12.22              | 14.91                | 0.010    | 20106.6                             |
| Z $\rightarrow$ eebb+Np1          | 109301  | 147838    | 4.95               | 6.04                 | 0.008    | 24480.5                             |
| Z $\rightarrow$ eebb+Np2          | 109302  | 39985     | 1.96               | 2.39                 | 0.012    | 16721.7                             |
| Z $\rightarrow$ eebb+Np3          | 109303  | 10000     | 0.95               | 1.16                 | 0.023    | 8628.1                              |
| Z $\rightarrow$ $\mu\mu$ bb+Np0   | 109305  | 299714    | 12.28              | 14.98                | 0.010    | 20005.5                             |
| Z $\rightarrow$ $\mu\mu$ bb+Np1   | 109306  | 144742    | 4.92               | 6.00                 | 0.008    | 24114.0                             |
| Z $\rightarrow$ $\mu\mu$ bb+Np2   | 109307  | 39952     | 1.92               | 2.34                 | 0.012    | 17056.0                             |
| Z $\rightarrow$ $\mu\mu$ bb+Np3   | 109308  | 10000     | 0.94               | 1.14                 | 0.023    | 8748.8                              |
| Z $\rightarrow$ $\tau\tau$ bb+Np0 | 109310  | 299664    | 12.37              | 15.09                | 0.010    | 19856.6                             |
| Z $\rightarrow$ $\tau\tau$ bb+Np1 | 109311  | 148999    | 4.88               | 5.95                 | 0.008    | 25026.7                             |
| Z $\rightarrow$ $\tau\tau$ bb+Np2 | 109312  | 39968     | 1.93               | 2.35                 | 0.012    | 16974.4                             |
| Z $\rightarrow$ $\tau\tau$ bb+Np3 | 109313  | 10000     | 0.97               | 1.19                 | 0.024    | 8434.6                              |

Table C.1: List of MC samples for 10 TeV analysis, SUSY signal and SM background. Cross sections from AMI database [126], K-factors for calculation of  $\sigma_{NLO}$  from [78, 170]

| Tag | $p_T$ range                 |
|-----|-----------------------------|
| J1  | 17 GeV $< p_T < 35$ GeV     |
| J2  | 35 GeV $< p_T < 70$ GeV     |
| J3  | 70 GeV $< p_T < 140$ GeV    |
| J4  | 140 GeV $< p_T < 280$ GeV   |
| J5  | 280 GeV $< p_T < 560$ GeV   |
| J6  | 560 GeV $< p_T < 1120$ GeV  |
| J7  | 1120 GeV $< p_T < 2240$ GeV |
| J8  | $p_T > 2240$ GeV            |

Table C.2: QCD di-jet  $p_T$  ranges corresponding to JX tags in MC sample names. [78]

| Dataset                 | # ev. | AOD   | $\sigma$ [pb] | Scale f. | $\int dt \mathcal{L}$ [fb $^{-1}$ ] |
|-------------------------|-------|-------|---------------|----------|-------------------------------------|
| directgauginos_p1       | 5000  | 18.98 | 0.76          | 263.4    |                                     |
| directgauginos_p2       | 5000  | 18.96 | 0.76          | 263.7    |                                     |
| directgauginos_p3       | 5000  | 3.25  | 0.13          | 1536.6   |                                     |
| directgauginos_p4       | 5000  | 1.89  | 0.08          | 2646.9   |                                     |
| directgauginos_sle_p01  | 5000  | 8.07  | 0.32          | 619.9    |                                     |
| directgauginos_sle_p02  | 5000  | 8.15  | 0.33          | 613.3    |                                     |
| directgauginos_sle_p03  | 5000  | 2.90  | 0.12          | 1722.9   |                                     |
| directgauginos_sle_p04  | 5000  | 2.88  | 0.12          | 1734.9   |                                     |
| directgauginos_sle_p05  | 5000  | 2.83  | 0.11          | 1768.0   |                                     |
| directgauginos_sle_p06  | 5000  | 2.84  | 0.11          | 1763.0   |                                     |
| directgauginos_sle_p07  | 5000  | 2.17  | 0.09          | 2308.4   |                                     |
| directgauginos_sle_p08  | 5000  | 4.58  | 0.18          | 1090.8   |                                     |
| directgauginos_sle_p09  | 5000  | 0.93  | 0.04          | 5376.9   |                                     |
| directgauginos_sle_p10  | 5000  | 3.84  | 0.15          | 1301.7   |                                     |
| directgauginos_sle_p11  | 5000  | 0.48  | 0.02          | 10491.0  |                                     |
| directgauginos_sle_p12  | 5000  | 3.67  | 0.15          | 1360.9   |                                     |
| directgauginos_stau_p01 | 5000  | 8.86  | 0.35          | 564.1    |                                     |
| directgauginos_stau_p02 | 5000  | 9.00  | 0.36          | 555.7    |                                     |
| directgauginos_stau_p03 | 5000  | 3.29  | 0.13          | 1519.3   |                                     |
| directgauginos_stau_p04 | 5000  | 3.24  | 0.13          | 1544.6   |                                     |
| directgauginos_stau_p05 | 5000  | 3.25  | 0.13          | 1540.8   |                                     |
| directgauginos_stau_p06 | 5000  | 3.11  | 0.12          | 1607.7   |                                     |
| directgauginos_stau_p07 | 5000  | 1.45  | 0.06          | 3445.9   |                                     |
| directgauginos_stau_p08 | 5000  | 1.33  | 0.05          | 3765.1   |                                     |
| directgauginos_stau_p09 | 5000  | 0.64  | 0.03          | 7804.0   |                                     |
| directgauginos_stau_p10 | 5000  | 0.57  | 0.02          | 8827.7   |                                     |
| directgauginos_stau_p11 | 5000  | 0.33  | 0.01          | 15046.6  |                                     |
| directgauginos_stau_p12 | 5000  | 0.27  | 0.01          | 18208.3  |                                     |
| gauginoshake_p1         | 10000 | 25.78 | 0.52          | 387.9    |                                     |
| gauginoshake_p2         | 5000  | 16.05 | 0.64          | 311.5    |                                     |
| gauginoshake_p3         | 5000  | 18.59 | 0.74          | 269.0    |                                     |

Table C.3: List of SUSY MSSM24 samples for 10 TeV analysis. [177]

# Appendix D

## List of outreach contributions

- **8. June 2004** The Venus transit - co-organized of the Department of physics booth at a general public event in Frognerparken.
- **Sep. 2004** CERN 50th anniversary - Co-organizing of the department of physics National Science Fair booth. Exhibition organized in collaboration with CERN.
- **Sep. 2004** UngForsk - (YoungResearch) - co-organizing experimental particle physics booth based on the CERN 50 years exhibition
- **2005-2008** International Mastreclass in particle Physics, organizer of the event in Oslo, see Section 13.1.
- **2006 - 2009 (Jan.)** Member of the Department of physics outreach committee
- **22. - 24. May 2007** Media excursion to CERN - participation and 20 min. plenary presentation of the ATLAS experiment.
- **May 2007** “Verdens største eksperiment” Article by NTB - Norwegian News Agency
- **10. Sep. 2008** Interview “Dagsnytt atten” - NRK-Radio - news program  
Short interview on TV-Norge news program  
Interviews related to the LHC start-up.
- **2004-2008** School visits, both at the university and at the schools
- **14. Feb. 2009** “Det må være noe mer” (There has to be something more) Radio lecture NRK-P2. Section 13.3
- **7. May 2009** “Angels and Demons” - meeting with the press before the film. “Fysikk og religion: Søken etter det grunnleggende” (Physics and religion: a search for the fundamental)
- **21. May 2009** “Fysikk og religion: Søken etter det grunnleggende” (Physics and religion: a search for the fundamental), feature areicle in [www.forskning.no](http://www.forskning.no)  
<http://www.forskning.no/artikler/2009/mai/220637>
- **12. Aug. 2009** “Små partikler og universets store mysterier” (Small particles and the great mysteries of the universe), popular science lecture for first year students.

- **16. Oct. 2009** “Verdt å vite” - NRK P2 - popular science radio program. Interview/comment about methods of wine dating with the help of proton beams.
- **18. - 19. Sep. 2009** Co-organized of the department booth at National Science Fair. The Event is described in Section 13.2
- **8. Jan. 2010** “Siste nytt om Large Hadron Collider ved CERN” (Latest news from the LCH experiments) Faglig pedagogisk dag 2010.  
<http://www.uio.no/publikum/skole/fagpeddag/presentasjoner/kp-fagped.080110.pdf>

# Appendix E

## List of papers

### ATLAS and Grid computing

- *Usage statistics and usage patterns on the NorduGrid: Analyzing the logging information collected on one of the largest production Grids of the world*, B. Kónya, P. A. M. Eerola, T. J. C. Ekelöf, M. Ellert, J. Hansen, A. Konstantinov, J. Nielsen, F. Ould-Saada, O. Smirnova, A. Wäänänen, U. Erkarslan, K. Pajchel (corresponding author), CHEP04 poster and proceedings
- *Performance of the NorduGrid ARC and the Dulcinea Executor in ATLAS Data Challenge 2*, P. A. M. Eerola, T. J. C. Ekelöf, O. Smirnova, M. Ellert, J. Hansen, A. Konstantinov, B. Kónya, J. Nielsen, F. Ould-Saada, A. Wäänänen, R. Sturrock, H. Jensen, J. Kleist, D. Kalici, A. Teras, H. Heller, J. Kennedy, G. Duckeck, T. Mycklebust, K. Pajchel, A. Read, H. Riiser, M. Hanshaugen, J. Sunde, A. Filipcic, M. Tadel, L. Nixon, J. Lindemann, L. Malinowsky, N. Smeds, A. Sandgren, M. Wadenstein, C. Haerberli, CHEP04 proceedings
- *An RPM-based kit for ATLAS production*, S. Ferrag, J. L. Nielsen, K. Pachjel, A. L. Read, F. Ould-Saada, ATL-COM-SOFT-2006-010
- *Data Management for the World's Largest Machine*, S. Haug, F. Ould-Saada, K. Pajchel, A. L. Read, *PARA* (2006) 480-488, conference proceedings
- *Complete distributed computing environment for a HEP experiment: Experience with ARC-connected infrastructure for ATLAS*, D. Cameron, F. Ould-Saada, K. Pajchel, A. Read, B. H. Samset, A. Taga, *J. Phys. Conf. Ser.* **119** (2008), conference proceedings CHEP07
- *The Advanced Resource Connector for Distributed LHC Computing*, D. Cameron, A. Konstantinov, F. Ould-Saada, Katarina Pajchela, Alexander Read, B. Samset, A. Taga, PoS(ACAT08)046, conference proceedings
- *Ganga: A tool for computational-task management and easy access to Grid resources*, J. T. Moscicki, F. Brochu, J. Ebke, U. Egede, J. Elmsheuser, K. Harrison, R. W. L. Jones, H. C. Lee, D. Liko, A. Maier, A. Muraru, G. N. Patrick, K. Pajchel, W. Reece, B. H. Samset, M. W. Slater, A. Soroko, C. L. Tan, D. C. van der Ster, M. Williams, *Computer Physics Communications* **180** (2009) no. 11 2302-2316, [arXiv:0902.2685v2](https://arxiv.org/abs/0902.2685v2)

- *Performance of an ARC-enabled computing grid for ATLAS/LHC physics analysis and Monte Carlo production under realistic conditions*, B. H. Samset, D. Cameron, M. Ellert, A. Filipcic, M. Gronager, J. Kleist, S. Maffioletti, F. Ould-Saada, K. Pajchel, A. L. Read, A. Taga and the ATLAS Collaboration, proceedings CHEP09, submitted to Journal of Physics: Conference Series, IOP Publishing.
- *Distributed Analysis in ATLAS using GANGA*, J. Elmsheuser, F. Brochu, G. Cowan, U. Egede, B. Gaidioz, H.-C. Lee, A. Maier, J. Moscicki, K. Pajchel, W. Reece, B. Samset, M. Slater, A. Soroko, D. van der Ster, M. Williams, proceedings CHEP09, submitted to Journal of Physics: Conference Series, IOP Publishing.

## ATLAS physics

- *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, The ATLAS Collaboration, CERN-OPEN-2008-020, arXiv:0901.0512v4, contributions to chapters “Prospects for Supersymmetry Discovery Based on Inclusive Searches” and “Multi-Lepton Supersymmetry Searches”
- *Prospects for SUSY discovery based on inclusive searches with the ATLAS detector at the LHC (Long Version)* J. Abdallah et. al., ATL-COM-PHYS-2009-261
- *Searching for new physics in events with three leptons with ATLAS* K. Pajchel, on behalf of the ATLAS collaboration, PoS(2008LHC)114, ATL-PHYS-PROC-2009-004, poster and proceedings Physics at LHC 2008, Split, Croatia
- *Searches for SUSY in trilepton final states at ATLAS*, K. Pajchel, L. Bugge, F. Ould-Saada, ATL-PHYS-INT-2009-024
- *Multi-lepton SUSY searches with the ATLAS detector* K. Pajchel, on behalf of the ATLAS collaboration, ATL-PHYS-SLIDE-2009-148 ; ATL-COM-PHYS-2009-273, presentation SUSY09, Boston, USA
- *Multi-lepton SUSY searches with the ATLAS detector*, K. Pajchel, on behalf of the ATLAS collaboration, ATL-PHYS-PROC-2009-095, proceedings SUSY09, Boston, USA, submitted to American Institute of Physics



# Bibliography

- [1] **K. Pajchel**, “P2 akademiet 14. february 2009, nrk, norwegian national broadcasting company.”  
<http://www.nrk.no/programmer/radio/p2-akademiet/1.6490798>.
- [2] **K. Pajchel**, “Små partikler og universets store mysterier.” Talk during the introductory week for starting students at the department, 12. August 2009,.
- [3] **K. Pajchel**, “Små partikler og universets store mysterier.” Talk during Fagpedagogisk Dag, UiO 8. January 2010,  
<http://www.uio.no/publikum/skole/fagpeddag/realfag.html>.
- [4] F. Halzen, A. D. Martin, *Quarks and Leptons*. Wiley, 1984.
- [5] M. E. Peskin, D. V. Schroeder, *An Introduction to Quantum Field Theory; 1995 ed.* Westview, 1995.
- [6] F. Mandl, G. Shaw, *Quantum Field Theory*. Wiley, 1984.
- [7] P. J. Mulders, “Introduction to quantum field theory.”  
<http://www.nat.vu.nl/~mulders/QFT-0.pdf>, 2009.
- [8] W. N. Cottingham, D. A. Greenwood, *An Introduction to the Standard Model of Particle Physics*. Cambridge, 2007.
- [9] M. Krämer, F. J. P. Soler, *Large Hadron Collider Phenomenology*. Institute of Physics Publishing, 2004.
- [10] C. D. Anderson, *The positive electron*, *Phys. Rev.* **43** (Mar, 1933) 491–494.
- [11] A. Salam, J. C. Ward, *Weak and electromagnetic interactions.*, *Nuovo Cimento, X. Ser.* **11** (1959) 568–577.
- [12] M. Gell-Mann, Y. Ne’eman, *The eightfold way*. Benjamin, 1964.
- [13] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, *Phys. Lett.* **12** (1964) 132–133.
- [14] P. W. Higgs, *Broken symmetries and the masses on the gauge bosons*, *Phys. Rev. Lett.* **13** (1964) 508–509.
- [15] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, *Phys. Rev.* **145** (1966) 1156–1163.

- [16] F. Englert, R. Brout, *Broken symmetries and the masses of the gauge vector bosons*, *Phys. Rev. Lett.* **13** (1964) 321–322.
- [17] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global conservation laws and massless particles*, *Phys. Rev. Lett.* **13** (1964) 585–587.
- [18] G. 't Hooft, M. J. G. Veltman,, *Regularization and Renormalization of Gauge Fields*, *Nucl. Phys.* **B44** (1972) 189–213.
- [19] S. L. Glashow, *Nucl. Phys.* **22** (1961) 579, S. Weinberg, *Phys. Rev. Lett.* **19** (1967) 1264, A. Salam, in *Elementary Particle Theory*, ed. N. Svartholm, Stockholm, “Almqvist and Wiksell” (1968), 367.
- [20] J. Goldstone, *Field Theories with Superconductor Solutions*, *Nuovo Cim.* **19** (1961) 154–164.
- [21] S. L. Glashow *Nucl. Phys.* **22** (1961) 579.
- [22] Y. Fukuda *et al.*, *Evidence for oscillation of atmospheric neutrinos*, *Phys. Rev. Lett.* **81** (Jul, 1998) 1562.
- [23] LEP Working Group for Higgs boson searches , R. Barate *et. al.*, *Search for the standard model Higgs boson at LEP*, *Phys. Lett.* **B565** (2003) 61–75, [[hep-ex/0306033v1](#)].
- [24] ALEPH Collaboration *et. al.*, *Precision Electroweak Measurements and Constraints on the Standard Model*, [arXiv:0911.2604](#).
- [25] CDF and D0 Collaborations , T. Aaltonen *et. al.*, *Combination of Tevatron searches for the standard model Higgs boson in the  $W^+W^-$  decay mode*, [arXiv:1001.4162](#).
- [26] The TEVNPH Working Group, *Combined CDF and D0 Upper Limits on Standard Model Higgs-Boson Production with up to 4.2 fb<sup>-1</sup> of Data*, [arXiv:0903.4001](#).
- [27] N. Cabibbo, *Unitary symmetry and leptonic decays*, *Phys. Rev. Lett.* **10** (Jun, 1963) 531–533.
- [28] M. Kobayashi, T. Maskawa, *CP-violation in the renormalizable theory of weak interaction*, *Progress of Theoretical Physics* **49** (1973), no. 2 652–657.
- [29] J. R. Ellis, *Limits of the standard model*, [hep-ph/0211168](#).
- [30] P. Renton, *Global Electroweak Fits and the Higgs Boson Mass*, [arXiv:0809.4566](#).
- [31] H. Flacher *et. al.*, *Gfitter - Revisiting the Global Electroweak Fit of the Standard Model and Beyond*, *Eur. Phys. J.* **C60** (2009) 543–583, [[arXiv:0811.0009](#)].
- [32] Particle Data Group , C. Amsler *et. al.*, *Review of particle physics*, *Phys. Lett.* **B667** (2008) 1.
- [33] J. R. Ellis, *Supersymmetry for Alp hikers*, [hep-ph/0203114](#).

- [34] A. Djouadi, *The Anatomy of electro-weak symmetry breaking. II. The Higgs bosons in the minimal supersymmetric model*, *Phys. Rep.* **459** (2008) 1–241, [[hep-ph/0503173](#)].
- [35] I. J. R. Aitchison, *Supersymmetry and the MSSM: An Elementary introduction*, [hep-ph/0505105](#).
- [36] Y. A. Golfand and E. P. Likhtman, *Extension of the Algebra of Poincare Group Generators and Violation of  $p$  Invariance*, *JETP Lett.* **13** (1971) 323–326.
- [37] J. R. Ellis, *Beyond the standard model for hill walkers*, [hep-ph/9812235](#).
- [38] W. de Boer, *Grand unified theories and supersymmetry in particle physics and cosmology*, *Prog. Part. Nucl. Phys.* **33** (1994) 201–302, [[hep-ph/9402266](#)].
- [39] M. E. Peskin, *Beyond the standard model*, [hep-ph/9705479](#).
- [40] S. P. Martin, *A Supersymmetry Primer*, [hep-ph/9709356](#).
- [41] H. E. Haber, R. Hempfling, *Can the mass of the lightest Higgs boson of the minimal supersymmetric model be larger than  $m(Z)$ ?*, *Phys. Rev. Lett.* **66** (1991) 1815–1818.
- [42] J. R. Ellis, G. Ridolfi, and F. Zwirner, *Radiative corrections to the masses of supersymmetric Higgs bosons*, *Phys. Lett.* **B257** (1991) 83–91.
- [43] Y. Okada, M. Yamaguchi, and T. Yanagida, *Upper bound of the lightest Higgs boson mass in the Minimal Supersymmetric Standard Model*, *Progress of Theoretical Physics* **85** (1991), no. 1 1–5.
- [44] J. Ellis, J. R. Espinosa, G. F. Giudice, A. Hoecker, and A. Riotto, *The Probable Fate of the Standard Model*, *Phys. Lett.* **B679** (2009) 369–375, [[arXiv:0906.0954](#)].
- [45] M. Drees, S. P. Martin, *Implications of SUSY model building*, [hep-ph/9504324](#).
- [46] N. Arkani-Hamed, S. Dimopoulos, G. F. Giudice, and A. Romanino, *Aspects of split supersymmetry*, *Nucl. Phys.* **B709** (2005) 3–46, [[hep-ph/0409232](#)].
- [47] J. L. Daz-Cruz, J. R. Ellis, K. A. Olive, and Y. Santoso, *On the feasibility of a stop nlsp in gravitino dark matter scenarios*, [oai:cds.cern.ch:1013671](#), *J. High Energy Phys.* **05** (Jan, 2007) 003. 31 p.
- [48] LEPSUSYWG, ALEPH, DELPHI, L3 and OPAL experiments.  
<http://lepsusy.web.cern.ch/lepsusy/>.
- [49] M. Battaglia *et. al.*, *Post-LEP CMSSM benchmarks for supersymmetry*, [hep-ph/0112013](#).
- [50] M. Battaglia *et. al.*, *Updated post-WMAP benchmarks for supersymmetry*, *Eur. Phys. J.* **C33** (2004) 273–296, [[hep-ph/0306219](#)].
- [51] H. Baer and C. Balazs,  $\chi^2$  analysis of the minimal supergravity model including WMAP,  $g_\mu$ -2 and  $b \rightarrow s$  gamma constraints, [hep-ph/0303114](#).

- [52] WMAP Collaboration , G. Hinshaw *et. al.*, *Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Data Processing, Sky Maps, & Basic Results, Astrophys. J. Suppl.* **180** (2009) 225–245, [arXiv:0803.0732].
- [53] Heavy Flavor Averaging Group , E. Barberio *et. al.*, *Averages of  $b$ -hadron and  $c$ -hadron Properties at the End of 2007*, arXiv:0808.1297.
- [54] Muon G-2 , G. W. Bennett *et. al.*, *Final report of the muon E821 anomalous magnetic moment measurement at BNL*, *Phys. Rev.* **D73** (2006) 072003, [hep-ex/0602035].
- [55] The CDMS-II , Z. Ahmed *et. al.*, *Results from the Final Exposure of the CDMS II Experiment*, arXiv:0912.3592.
- [56] H. V. Klapdor-Kleingrothaus, *Search for dark matter by GENIUS-TF and GENIUS*, *Nucl. Phys. Proc. Suppl.* **110** (2002) 58–60, [hep-ph/0206250].
- [57] CDF Collaboration , T. Adams, *SUSY Searches at the Tevatron*, arXiv:0808.0728.
- [58] CDF Collaboration, *Update of the unified trilepton search with  $3.2\text{ fb}^{-1}$  of data*, 2009. [http://www-cdf.fnal.gov/physics/exotic/r2a/20090521.trilepton\\_3fb/Welcome.html](http://www-cdf.fnal.gov/physics/exotic/r2a/20090521.trilepton_3fb/Welcome.html).
- [59] D0 Collaboration , V. M. Abazov *et. al.*, *Search for associated production of charginos and neutralinos in the trilepton final state using  $2.3\text{ fb}^{-1}$  of data*, *Phys. Lett.* **B680** (2009) 34–43, [arXiv:0901.0646].
- [60] H. Baer, X. Tata, *Weak Scale Supersymmetry*. Cambridge, 2006.
- [61] A. Cesarini, F. Fucito, A. Lionetto, A. Morselli, and P. Ullio, *The galactic center as a dark matter gamma-ray source*, *Astropart. Phys.* **21** (2004) 267–285, [astro-ph/0305075].
- [62] A. W. Strong, I. V. Moskalenko, and O. Reimer, *Diffuse Galactic continuum gamma rays. A model compatible with EGRET data and cosmic-ray measurements*, *Astrophys. J.* **613** (2004) 962–976, [astro-ph/0406254].
- [63] D. Hooper, G. Zaharijas, D. P. Finkbeiner, and G. Dobler, *Prospects For Detecting Dark Matter With GLAST In Light Of The WMAP Haze*, *Phys. Rev.* **D77** (2008) 043511, [arXiv:0709.3114].
- [64] O. Adriani *et. al.*, *A new measurement of the antiproton-to-proton flux ratio up to 100 GeV in the cosmic radiation*, *Phys. Rev. Lett.* **102** (2009) 051101, [arXiv:0810.4994].
- [65] K. Abe *et. al.*, *Measurement of cosmic-ray low-energy antiproton spectrum with the first BESS-Polar Antarctic flight*, *Phys. Lett.* **B670** (2008) 103–108, [arXiv:0805.1754].
- [66] A. Morselli and I. V. Moskalenko, *Status of indirect searches in the PAMELA and Fermi era*, arXiv:0811.3526.

- 
- [67] PAMELA Collaboration , O. Adriani *et. al.*, *Observation of an anomalous positron abundance in the cosmic radiation*, *Nature* **458** (2009) 607–609, [[arXiv:0810.4995](#)].
- [68] J. Chang *et. al.*, *An excess of cosmic ray electrons at energies of 300–800 GeV*, *Nature* **456** (2008) 362–365.
- [69] D. Hooper, P. Blasi, and P. D. Serpico, *Pulsars as the Sources of High Energy Cosmic Ray Positrons*, *JCAP* **0901** (2009) 025, [[arXiv:0810.1527](#)].
- [70] O. Bruning, P. Collier, *Building a behemoth*, *Nature* **448** (2007) 285–289.
- [71] *Interim summary report on the analysis of the 19 September 2008 incident at the LHC*, October, 2008. CERN/AT/PhL, Document EDMS 973073.
- [72] S. Stapnes, *Detector challenges at the LHC*, *Nature* **448** (2007) 290–296.
- [73] ATLAS Collaboration , G. Aad *et. al.*, *The ATLAS experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.
- [74] Electron identification.  
<https://twiki.cern.ch/twiki/bin/view/Main/ElectronIDIsEM>.
- [75] D. Attree *et. al.*, *The evaporative cooling system for the ATLAS inner detector*, *JINST* **3** (2008) P07003.
- [76] S. Myers and F. Zimmermann, *Chamonix 2009 workshop on LHC performance*, February, 2009. CERN-ATS-2009-001.
- [77] LHC Performance Workshop - Chamonix 2010, January, 2010. LHC Performance Workshop.
- [78] The ATLAS Collaboration , G. Aad *et. al.*, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, [arXiv:0901.0512](#).
- [79] UCLA Press Release. <http://www.lk.cs.ucla.edu/LK/Bib/REPORT/press.html> .
- [80] I. Foster and C. Kesselman, eds., *The grid: blueprint for a new computing infrastructure*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1999.
- [81] I. Foster, C. Kesselman, and S. Tuecke, *The anatomy of the grid: Enabling scalable virtual organizations*, *International Journal of Supercomputer Applications* **15** (2001), no. 3.
- [82] I. Foster, *What is the grid? - a three point checklist*, *GRIDtoday* **1** (July, 2002).
- [83] R. Alfieri *et. al.*, *From gridmap-file to VOMS: managing authorization in a Grid environment*, *Future Gener. Comput. Syst.* **21** (2005), no. 4 549–558.
- [84] *ATLAS computing: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 2005. ATLAS-TDR-017; CERN-LHCC-2005-022.
- [85] Open Science Grid, *A blueprint for the open science grid*, 2009.  
<http://www.opensciencegrid.org/documents/>.

- [86] Enabling Grids for E-sciencE (EGEE). <http://www.eu-egee.org/>.
- [87] NorduGrid. <http://www.nordugrid.org/>.
- [88] Nordic DataGrid Facility. <http://www.ndgf.org>.
- [89] L. Fischer, M. Grønager, J. Kleist, and O. Smirnova, *A distributed Tier-1*, *J. Phys. Conf. Ser.* **119** (2008) 052016.
- [90] R. Brun, F. Rademakers, *ROOT – an object oriented data analysis framework*, *Nuclear Instruments and Methods in Physics Research Section A* **389** (1997), no. 1-2 81 – 86. New Computing Techniques in Physics Research.
- [91] R. Jones, “The atlas computing model.” School on HEP@TR-GRID, 30 April - 2 May 2008, Ankara, Turkey.
- [92] ATLAS DDM.  
<https://twiki.cern.ch/twiki/bin/view/Atlas/DistributedDataManagement>.
- [93] S. Albrand *et. al.*, *Atlas dataset nomenclature*, Tech. Rep. ATL-COM-GEN-2007-003, CERN, Geneva, Nov., 2007.
- [94] M. Lassnig *et. al.*, *Managing ATLAS data on a petabyte-scale with DQ2*, in *Journal of Physics: Conference Series*, (Bristol, England), Institute of Physics Publishing, September.
- [95] S. Traylen, *LCG VOBx operations recommendations and questionnaire*, Tech. Rep. CERN-LCG-EDMS-655277, LCG Operations Workshop, December, 2005.
- [96] LCG File Catalog <https://twiki.cern.ch/twiki/bin/view/EGEE/GliteLFC> .
- [97] Storage Resource Management (SRM) Working Group.  
<https://sdm.lbl.gov/srm-wg/> .
- [98] ATLAS Dashboard. <http://dashboard.cern.ch/atlas/>.
- [99] PanDA production monitor.  
<http://panda.cern.ch:25880/server/pandamon/query?dash=prod>.
- [100] T. Maeno, *PanDA: distributed production and distributed analysis system for ATLAS*, *Journal of Physics: Conference Series* **119** (2008), no. 6 062036 (4pp).
- [101] PanDA distributed analysis.  
<https://twiki.cern.ch/twiki/bin/view/Atlas/PanDA>.
- [102] J. T. Moscicki, F. Brochu, J. Ebke, U. Egede, J. Elmsheuser, K. Harrison, R. W. L. Jones, H. C. Lee, D. Liko, A. Maier, A. Muraru, G. N. Patrick, **K. Pajchel**, W. Reece, B. H. Samset, M. W. Slater, A. Soroko, C. L. Tan, D. C. van der Ster, and M. Williams, *Ganga: A tool for computational-task management and easy access to grid resources.*, *Computer Physics Communications* **180** (2009), no. 11 2303–2316. [arXiv:0902.2685v2](https://arxiv.org/abs/0902.2685v2).
- [103] E. Laure *et. al.*, *Programming the Grid with gLite*, Tech. Rep. EGEE-TR-2006-001, CERN, Geneva, Mar, 2006.

- [104] gLite, Lightweight Middleware for Grid Computing  
<http://glite.web.cern.ch/glite/> .
- [105] M. Ellert *et. al.*, *Advanced resource connector middleware for lightweight computational Grids, Future Gener. Comput. Syst.* **23** (2007), no. 2 219–240.
- [106] O.Smirnova, *Extended resource specification language*, Tech. Rep. NORDUGRID-MANUAL-4, The NorduGrid Collaboration, October, 2009.  
<http://www.nordugrid.org/documents/xrsl.pdf>.
- [107] O. Smirnova, *The Grid Monitor, usage manual*, Tech. Rep. NORDUGRID-MANUAL-5, The NorduGrid Collaboration, September, 2003.  
<http://www.nordugrid.org/documents/monitor.pdf>.
- [108] I. T. Foster, *Globus toolkit version 4: Software for service-oriented systems.*, in *NPC* (H. Jin, D. A. Reed, and W. Jiang, eds.), vol. 3779 of *Lecture Notes in Computer Science*, pp. 2–13, Springer, 2005.
- [109] A. Konstantinov, *The nordugrid grid manager and gridftp server: Description and administrator's manual*, Tech. Rep. NORDUGRID-TECH-2, The NorduGrid Collaboration, January, 2008. <http://www.nordugrid.org/documents/GM.pdf>.
- [110] P. A. M. Eerola, T. J. C. Ekelöf, O. Smirnova, M. Ellert, J. Hansen, A. Konstantinov, B. Kónya, J. Nielsen, F. Ould-Saada, A. Wäänänen, R. Sturrock, H. Jensen, J. Kleist, D. Kalici, A. Teras, H. Heller, J. Kennedy, G. Duckeck, J. Mycklebust, **Pajchel, K**, A. Read, H. Riiser, M. Hanshaugen, S. Sunde, A. Filipcic, M. Tadel, L. Nixon, J. Lindemann, L. Malinowsky, N. Smeds, A. Sandgren, M. Wadenstein, and C. Haeberli, *Performance of the nordugrid arc and the dulcinea executor in atlas data challenge 2, CHEP 04, proceedings* (2005).
- [111] **Pajchel, K**, B. Kónya, P. A. M. Eerola, T. J. C. Ekelöf, M. Ellert, J. Hansen, A. Konstantinov, J. Nielsen, F. Ould-Saada, O. Smirnova, A. Wäänänen, and U. Erkarlsan, *Usage statistics and usage patterns on the nordugrid analysing the logging information on one of the largest production grids in the world, CHEP 04, proceedings* (2005).
- [112] MySQL open source database. <http://www.mysql.com/>.
- [113] PHP Hypertext Preprocessor. <http://php.net/index.php>.
- [114] JpGraph Object-Oriented Graph creating library for PHP.  
<http://www.aditus.nu/jpgraph/>.
- [115] GD Library graphical library. <http://www.libgd.org/> .
- [116] A.Konstantinov, *The logger service*, Tech. Rep. NORDUGRID-TECH-11, The NorduGrid Collaboration, September, 2009.  
<http://www.nordugrid.org/documents/Logger.pdf>.
- [117] S. Ferrag, J. L. Nielsen, **Pachjel, K**, A. L. Read, and F. Ould-Saada, *An rpm-based kit for atlas production*, Tech. Rep. ATL-COM-SOFT-2006-010, CERN, Geneva, Jul, 2005.

- [118] Red Hat Package Manager RPM. <http://www.rpm.org/>.
- [119] C. Arnault, *CMT: A software configuration management tool*, . Prepared for International Conference on Computing in High- Energy Physics and Nuclear Physics (CHEP 2000), Padova, Italy, 7-11 Feb 2000.
- [120] Configuration Management Tool (CMT). <http://www.cmtsite.org/>.
- [121] Pacman. <http://physics.bu.edu/pacman/>.
- [122] Python Programming Language – Official Website. <http://www.python.org/>.
- [123] M. Dobbs and J. B. Hansen, *The HepMC C++ Monte Carlo event record for High Energy Physics*, *Comput. Phys. Commun.* **134** (2001) 41–46.
- [124] S. Agostinelli *et. al.*, *GEANT4: A simulation toolkit*, *Nucl. Instrum. Meth.* **A506** (2003) 250–303.
- [125] E. Richter-Was, D. Froidevaux, and L. Poggioli, *Atlfast 2.0 a fast simulation package for atlas*, Tech. Rep. ATL-PHYS-98-131, CERN, Geneva, Nov, 1998.
- [126] ATLAS Metadata Interface - AMI.  
<http://ami.in2p3.fr/opencms/opencms/AMI/www/>.
- [127] S. Albrand, T. Doherty, J. Fulachier, and F. Lambert, *The ATLAS Metadata Interface*, *CHEP 07, proceedings* (2007).  
<http://hal.in2p3.fr/in2p3-00192624/fr/>.
- [128] Kennedy J., ATLAS Production System Monitoring, CHEP07.  
[http://www.etp.physik.uni-muenchen.de/dokumente/talks/chep07\\_jkennedy.pdf](http://www.etp.physik.uni-muenchen.de/dokumente/talks/chep07_jkennedy.pdf) .
- [129] ATLAS Eowyn Supervisor.  
<https://twiki.cern.ch/twiki/bin/view/Atlas/EowynSupervisor>.
- [130] J.Nielsen, *ARClib, a client library for ARC*, Tech. Rep. NORDUGRID-TECH-12, The NorduGrid Collaboration, February, 2006.  
<http://www.nordugrid.org/documents/arclib.pdf>.
- [131] NorduGrid production executor Dulcinea and tools  
<http://atlas-sw.cern.ch/cgi-bin/viewcvs-atlas.cgi/offline/Production/Executors/ng/>.
- [132] G. Behrmann, D. Cameron, M. Ellert, J. Kleist, and A. Taga, *ATLAS DDM integration in ARC*, *J. Phys. Conf. Ser.* **119** (2008) 062015.
- [133] G. Behrmann, P. Fuhrmann, M. Grønager, and J. Kleist, *A distributed storage system with dCache*, *J. Phys. Conf. Ser.* **119** (2008) 062014.
- [134] LCG Savannah, Validation. <https://savannah.cern.ch/bugs/?group=validation> .
- [135] Valgrind <http://www.valgrind.org/>.
- [136] S. Haug, F. Ould-Saada, **K. Pajchel**, and A. L. Read, *Data management for the world's largest machine*, in *PARA*, pp. 480–488, 2006.



- 
- [137] D. Cameron, F. Outld-Saada, **K. Pajchel**, A. Read, B. H. Samset, and A. Taga, *Complete distributed computing environment for a HEP experiment: Experience with ARC-connected infrastructure for ATLAS*, *J. Phys. Conf. Ser.* **119** (2008) 062041.
- [138] M. Battaglia *et. al.*, *Proposed post-LEP benchmarks for supersymmetry*, *Eur. Phys. J.* **C22** (2001) 535–561, [[hep-ph/0106204](#)].
- [139] H. Baer, F. E. Paige, S. D. Protopopescu, and X. Tata, *Simulating supersymmetry with ISAJET 7.0 / ISASUSY 1.0*, [hep-ph/9305342](#).
- [140] F. Paige, S. Protopopescu, H. Baer, and X. Tata, *ISAJET 7.69: A Monte Carlo event generator for  $p p$ , anti- $p p$ , and  $e^+e^-$  reactions*, 2003.
- [141] G. Corcella *et. al.*, *HERWIG 6: An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, *JHEP* **01** (2001) 010, [[hep-ph/0011363](#)].
- [142] J. F. J. Butterworth and M. Seymour, *Multiparton interactions in photoproduction at hera*, *Z. Phys.* **C72** (1996) 637–646, [[hep-ph/9601371](#)].
- [143] M. S. W. Beenakker, R. Hopker and P. Zerwas, *Squark and gluino production at hadron colliders*, *Nucl. Phys.* **B492** (1997) 51–103, [[hep-ph/9610490](#)].
- [144] D. Stump *et. al.*, *Inclusive jet production, parton distributions, and the search for new physics*, *JHEP* **10** (2003) 046, [[hep-ph/0303013](#)].
- [145] A. Barr, Z plus b-jets.  
<https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ZplusBjet>.
- [146] W. Beenakker, R. Hopker, and M. Spira, *PROSPINO: A program for the PROduction of Supersymmetric Particles In Next-to-leading Order QCD*, [hep-ph/9611232](#).
- [147] G. Corcella *et. al.*, *HERWIG 6.5 release note*, [hep-ph/0210213](#).
- [148] S. Frixione and B. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **06** (2002) 029, [[hep-ph/0204244](#)].
- [149] S. Frixione, P. Naso, and B. Webber, *Matching NLO QCD and parton showers in heavy flavour production*, *JHEP* **08** (2003) 007, [[hep-ph/0305252](#)].
- [150] A. Barr, Gauge boson pairs with HERWIG.  
<https://twiki.cern.ch/twiki/bin/view/AtlasProtected/GaugeBosonPairs>.
- [151] B. P. Kersevan and E. Richter-Was, *The Monte Carlo event generator AcerMC version 2.0 with interfaces to PYTHIA 6.2 and HERWIG 6.5*, [hep-ph/0405247](#).
- [152] T. Sjostrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, *JHEP* **05** (2006) 026, [[hep-ph/0603175](#)].

- [153] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, *JHEP* **07** (2003) 001, [[hep-ph/0206293](#)].
- [154] S. Jadach, Z. Wąs, R. Decker, and J. H. Kühn, *The  $\tau$  decay library tauola, version 2.4*, *Computer Physics Communications* **76** (August, 1993) 361–380.
- [155] E. Barberio and Z. Was, *PHOTOS: A Universal Monte Carlo for QED radiative corrections. Version 2.0*, *Comput. Phys. Commun.* **79** (1994) 291–308.
- [156] Atlfast II <https://twiki.cern.ch/twiki/bin/view/Atlas/AtlfastII>.
- [157] K. Edmonds, S. Fleischmann, T. Lenz, C. Magass, J. Mechnich, and A. Salzburger, *The fast atlas track simulation (fatras)*, Tech. Rep. ATL-SOFT-PUB-2008-001. ATL-COM-SOFT-2008-002, CERN, Geneva, Mar, 2008.
- [158] FastCaloSim. <https://twiki.cern.ch/twiki/bin/view/Atlas/FastCaloSim>.
- [159] A. S. K. Cranmer, A. Farbin, *EventView - the design behind an analysis framework*, Tech. Rep. ATL-SOFT-PUB-2007-008. ATL-COM-SOFT-2007-012, CERN, Geneva, Sep, 2007.
- [160] A. Shibata, *TopView: An AOD analysis package for ATLAS top physics analysis*, 2007. ATL-COM-SOFT-2007-006.
- [161] ATLAS Muon Reconstruction . <https://twiki.cern.ch/twiki/bin/view/Atlas/MuonRecoPedia>.
- [162] D. Adams *et. al.*, *Track reconstruction in the ATLAS Muon Spectrometer with MOORE 007*, Tech. Rep. ATL-SOFT-2003-007, CERN, Geneva, May, 2003. revised version number 1 submitted on 2003-10-13 13:51:39.
- [163] S. Hassani *et. al.*, *A muon identification and combined reconstruction procedure for the ATLAS detector at the LHC using the (MUONBOY, STACO, MuTag) reconstruction packages*, *Nucl. Instrum. Meth.* **A572** (2007) 77–79.
- [164] O. Linossier, L. Poggioli,  *$h^0 \rightarrow zz^* \rightarrow 4l$  channel, in atlas signal reconstruction and reducible background rejection*, 1997. ATL-PHYS-97-101.
- [165] J. T. Linnemann, *Measures of significance in hep and astrophysics*, 2003. [physics/0312059](#).
- [166] A. De Santo, C. Potter, J. Dragic, *Trilepton SUSY signatures at ATLAS*, May, 2008. ATL-PHYS-INT-2009-024.
- [167] O. Brandt, A.J. Barr, P. Bruckman de Rentstrom, *Search strategies for Supersymmetry in tri-lepton final states with the ATLAS detector at the Large Hadron Collider*, Jan, 2008. ATL-COM-PHYS-2008-005.
- [168] M. Mangano *et. al.*, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, *JHEP* **07** (2003) 001, [[hep-ph/0206293](#)].

- 
- [169] E. L. Berger, Z. Sullivan, *Trilepton production at the CERN LHC: Standard model sources and beyond*, *Phys. Rev.* **D78** (2008) 034030, [[arXiv:0805.3720](https://arxiv.org/abs/0805.3720)].
  - [170] ATLAS Top working group, MC cross sections.  
<https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopMC2009>.
  - [171] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-kt jet clustering algorithm*, *Journal of High Energy Physics* **2008** (2008), no. 04 063.
  - [172] B. Abbott *et al.*, *Study on reconstructed object definition and selection for top physics*, Tech. Rep. ATL-COM-PHYS-2009-633, CERN, Geneva, Dec, 2009.
  - [173] S. D. Ellis, D. E. Soper, *Successive combination jet algorithm for hadron collisions*, *Phys. Rev. D* **48** (Oct, 1993) 3160–3166.
  - [174] G. Cowan, E. Gross, *Discovery significance with statistical uncertainty in the background estimate*, 2008. ATLAS Statistics Forum.
  - [175] J. M. Campbell, J. W. Huston, and W. J. Stirling, *Hard Interactions of Quarks and Gluons: A Primer for LHC Physics*, *Rept. Prog. Phys.* **70** (2007) 89, [[hep-ph/0611148](https://arxiv.org/abs/hep-ph/0611148)].
  - [176] E. A. Baltz, M. Battaglia, M. E. Peskin, and T. Wizansky, *Determination of dark matter properties at high-energy colliders*, *Phys. Rev. D* **74** (Nov, 2006) 103521.
  - [177] B. K. Gjelsten, “Detailed listing of generator information for mssm24 samples.”  
<http://folk.uio.no/borgeg/mssm24/PHENOGRIID/derived/>.
  - [178] Reprocessing of 900 GeV ATLAS data.  
<https://twiki.cern.ch/twiki/bin/view/Atlas/DataPreparationReprocessingPlanning>.
  - [179] H. Okawa *et al.*, *Calorimeter commissioning with cosmic rays: Cells, clusters, jets and missing ET*, Tech. Rep. ATL-COM-CAL-2008-007, CERN, Geneva, Dec, 2008.
  - [180] Presentations during ATLAS SUSY Working Group meetings, and for example *Discovery Potential and Exclusion Limits for R-Parity Conserving SUSY Signals at  $\sqrt{s} = 10$  TeV in ATLAS*, J. Poveda *et al.*, ATL-PHYS-INT-2009-114.
  - [181] **Pajchel, K.**, L. Bugge, and F. Ould-Saada, *Searches for susy in trilepton final states at atlas*, Tech. Rep. ATL-PHYS-INT-2009-024, CERN, Geneva, Nov, 2008.
  - [182] Kunnskapsdepartementet - Department of Education, “Tilstandsrapport for høyere utdanningsinstitusjoner 2009 - Survey of Norwegian educational institutions.”  
[http://media.regjeringen.no/kd/video/Tilstandsrapport\\_UH.09.1.pdf](http://media.regjeringen.no/kd/video/Tilstandsrapport_UH.09.1.pdf).
  - [183] E. K. Henriksen, C. Schreiner, and J. Sjaastad, “Kan realfagene surfe på yngrebølgen?.” <http://www.forskning.no/artikler/2010/januar/240555>.
  - [184] Kunnskapsdepartementet - Department of education, “Klima for forskning st. meld. nr. 30 (2008-2009) - climate for research white paper on research.”  
[http://www.regjeringen.no/upload/KD/Vedlegg/Forskning/climate\\_for\\_research.final.pdf](http://www.regjeringen.no/upload/KD/Vedlegg/Forskning/climate_for_research.final.pdf).
  - [185] World Year of Physics 2005, Web page:. <http://www.wyp2005.org/>.

- [186] The European Particle Physics Outreach Group, web page.  
<http://eppog.web.cern.ch/eppog/>.
- [187] Hands-on-CERN project page. <http://hands-on-cern.physto.se/>.
- [188] Stockholm House of Science. <http://www.houseofscience.se/>.
- [189] Hands-on-CERN in Norwegian web page.  
[http://www.fys.uio.no/epf/epf\\_outreach/hoc/hoc\\_v21sv/index.html](http://www.fys.uio.no/epf/epf_outreach/hoc/hoc_v21sv/index.html).
- [190] CloudchambersDotCom, web page. <http://www.cloudchambers.com/>.
- [191] J. P. Wellisch, C. Williams, and S. Ashby, *SCRAM: Software configuration and management for the LHC computing Grid project*, CoRR **cs.OH/0306014** (2003).
- [192] G. Corti, G. Barrand, I. Belyaev, M. Cattaneo, P. Charpentier, M. Frank, P. Koppenburg, P. Mato-Vila, F. Ranjard, and S. Roiser, *Software for the LHCb experiment*, *IEEE Trans. Nucl. Sci.* **53** (2006) 1323–1328.

# List of Figures

|      |   |     |
|------|---|-----|
| 1.1  | Meson nonet and baryon octet. . . . .   | 8   |
| 1.2  | The standard Model of particle physics . . . . .                                  | 10  |
| 1.3  | Overview of the LHC accelerator. . . . .  | 12  |
| 2.1  | Higgs potential. . . . .  | 20  |
| 2.2  | Tevatron Higgs exclusion plot and LEP EW Working Group “blue band” plot . . . . . | 26  |
| 3.1  | Loop diagrams contributing to the Higgs mass. . . . .                             | 30  |
| 3.2  | GUT unification of the coupling constants. . . . .                                | 40  |
| 3.3  | Evolution of the soft SUSY breaking masses. . . . .                               | 40  |
| 3.4  | Plot of $\chi^2/dof$ for the mSUGRA model . . . . .                               | 46  |
| 4.1  | Cross section of an LHC magnet. . . . .   | 48  |
| 4.2  | The CERN accelerator complex . . . . .  | 49  |
| 4.3  | The ATLAS detector . . . . .  | 52  |
| 4.4  | The ATLAS Inner Detector . . . . .  | 53  |
| 4.5  | The ATLAS calorimeter system . . . . .  | 55  |
| 4.6  | The ATLAS Muon spectrometer . . . . .   | 57  |
| 4.7  | The ATLAS Trigger and Data Acquisition System . . . . .                           | 59  |
| 4.8  | Setup for heater pulse tests. . . . .   | 60  |
| 4.9  | Combined track during M6 cosmic run . . . . .                                     | 61  |
| 5.1  | Geographical location of the NorduGrid ATLAS resources. . . . .                   | 72  |
| 5.2  | ARC job flow and status schema . . . . .  | 74  |
| 5.3  | NorduGrid - NDGF end-to-end Grid computing environment . . . . .                  | 75  |
| 5.4  | NGLogger DC2 overview. Example of a “project page”. . . . .                       | 78  |
| 5.5  | NGLogger example of user statistics . . . . .                                     | 79  |
| 5.6  | MC managed production work flow. . . . .  | 84  |
| 5.7  | ATLAS Production System . . . . .   | 87  |
| 5.8  | NorduGrid production system with the Duclinea executor. . . . .                   | 88  |
| 5.9  | NorduGrid Production System errors October 2006 . . . . .                         | 96  |
| 5.10 | NorduGrid Production System errors September 2007 . . . . .                       | 97  |
| 5.11 | NorduGrid Production System errors November, December 2008 . . . . .              | 98  |
| 5.12 | Number of DC2 and Rome production jobs, 2004-2005. . . . .                        | 100 |
| 5.13 | NorduGrid production highlights 2007 . . . . .                                    | 101 |
| 5.14 | Successful NDGF Tier-1 Christmas production 2007/2008 . . . . .                   | 102 |
| 5.15 | NDGF 03.15 – 07.01 2008 . . . . .   | 103 |
| 5.16 | NDGF biggest cloud. 02.07.2009 . . . . .  | 103 |
| 5.17 | NDGF production 01.01 – 12.31.2008 . . . . .                                      | 104 |

|      |   |     |
|------|---|-----|
| 5.18 | Ganga architecture . . . . .  | 105 |
| 5.19 | Ganga component classes . . . . .   | 107 |
| 6.1  | Hadron collider event. . . . .  | 115 |
| 6.2  | Feynman diagrams for direct gaigino production. . . . .   | 116 |
| 6.3  | Feynman diagrams for direct of $\chi_i^\pm \chi_j^\mp$ and $\chi_i^0 \chi_j^0$ pairs. . . . .   | 117 |
| 6.4  | Feynman diagrams for gluino production. . . . .   | 117 |
| 6.5  | Feynman diagrams for squark pair production via gluon scattering. . . . .   | 118 |
| 6.6  | Feynman diagrams for squark pair production via quark scattering. . . . .   | 118 |
| 6.7  | Feynman diagrams for gluino-squark production. . . . .  | 118 |
| 6.8  | Squark and gluino cross sections. . . . .   | 118 |
| 6.9  | Typical decay patterns for two common benchmark scenarios. . . . .  | 119 |
| 6.10 | The mSUGRA parameter space spanned by $m_{1/2}$ and $m_0$ . . . . .   | 123 |
| 6.11 | SU2 sparticle mass spectrum and Feynman diagram representing the dominant production process. . . . .   | 125 |
| 6.12 | SU3 sparticle mass spectrum and Feynman diagram representing the dominant production process. . . . .   | 126 |
| 6.13 | SU4 sparticle mass spectrum. . . . .  | 126 |
| 6.14 | Feynman diagrams for SM backgrounds: $t\bar{t}$ , $Zb$ , $WW$ . . . . .   | 127 |
| 6.15 | Feynman diagrams for SM backgrounds: $ZZ$ , $Z\gamma$ , $WZ$ . . . . .  | 127 |
| 6.16 | Schematic picture of the event simulation. . . . .  | 129 |
| 7.1  | Schematic overview of the different classes of leptons and their labels. . . . .  | 134 |
| 7.2  | $p_T$ distributions for muons from SU3 and $t\bar{t}$ . . . . .   | 135 |
| 7.3  | Calorimeter-based isolation $E_{\text{cal}}^{\Delta R=0.X}$ for muons in the signal sample SU3 and background $t\bar{t}$ . . . . .  | 136 |
| 7.4  | Track-based isolation $p_{T\Sigma\text{track}}^{\Delta R=0.X}$ for muons in the signal sample SU3 and background $t\bar{t}$ . . . . .   | 137 |
| 7.5  | Relative track-based isolation $p_{T\Sigma\text{track}}^{\Delta R=0.X}$ for muons in the signal sample SU3 and background $t\bar{t}$ . . . . .  | 137 |
| 7.6  | Profile plot of $E_{\text{cal}}^{\Delta R=0.3}$ and $p_{T\Sigma\text{track}}^{\Delta R=0.3}$ as function of $p_T$ for prompt muons in SU3 and for <i>extra</i> muons in the $t\bar{t}$ background sample. . . . . | 138 |
| 7.7  | Profile plot of $E_{\text{cal}}^{\Delta R=0.3}$ and $p_{T\Sigma\text{track}}^{\Delta R=0.3}$ as function of $p_T$ for prompt muons in SU3 and for <i>extra</i> muons in the $t\bar{t}$ background sample. . . . . | 138 |
| 7.8  | $\Delta R$ between muons and jets (jets have been through overlap removal with electrons). Histograms show muons in SU3 and $t\bar{t}$ samples. . . . .   | 139 |
| 7.9  | Profile of $\Delta R$ angular distance to closest jet for isolated prompt muons in SU3 and isolated <i>extra</i> muons in $t\bar{t}$ . . . . .  | 140 |
| 7.10 | Scatter plot of the transverse impact parameter significance $d0/\sigma(d0)$ vs. relative track-based isolation and $p_T$ along the x-axis vs. $p_{T\Sigma\text{track}}^{\Delta R=0.2}$ . . . . .                 | 141 |
| 7.11 | Track-based isolation after applying the $f_{40}(p_T)$ -cut for prompt muons in the signal sample SU3 and muons from jet in $t\bar{t}$ . . . . .  | 142 |
| 7.12 | Stack plot of $p_T$ distributions for electrons from SU3 and $t\bar{t}$ . . . . .   | 144 |
| 7.13 | Calorimeter-based isolation $E_{\text{cal}}^{\Delta R=0.X}$ for electrons in the signal sample SU3 and $t\bar{t}$ background sample. . . . .  | 145 |
| 7.14 | Track-based isolation $p_{T\Sigma\text{track}}^{\Delta R=0.X}$ for electrons in the signal sample SU3 and background $t\bar{t}$ . . . . .   | 146 |

|      |   |     |
|------|---|-----|
| 7.15 | Profile plot of $E_{\text{cal}}^{\Delta R=0.2}$ and $p_{T\Sigma\text{track}}^{\Delta R=0.2}$ as function of $p_T$ for prompt electrons in SU3 and <i>extra</i> electrons in the $t\bar{t}$ background sample. . . . .             | 147 |
| 7.16 | Relative track-based isolation $p_{T\Sigma\text{track}}^{\Delta R=0.X}$ for electrons in the signal sample SU3 and background $t\bar{t}$ . . . . .  | 147 |
| 7.17 | Profile plot of relative $E_{\text{cal}}^{\Delta R=0.2}$ and relative $p_{T\Sigma\text{track}}^{\Delta R=0.2}$ as function of $p_T$ for prompt electrons in SU3 <i>extra</i> electrons in the $t\bar{t}$ background sample. . . . | 148 |
| 7.18 | $\Delta R$ between electrons and jet objects without overlap removal and with overlap removal. . . . .  | 149 |
| 7.19 | Scatter plot of the relative track-based isolation $p_{T\Sigma\text{track}}^{\Delta R=0.2}$ along the y-axis vs. $p_T$ for prompt electrons in SU3 and <i>extra</i> electrons in $t\bar{t}$ . . . . .                             | 150 |
| 7.20 | Isolation cut optimisation for trilepton analysis, calorimeter-based and track-based isolation variables. . . . .   | 151 |
| 7.21 | Isolation cut optimisation for trilepton analysis, relative calorimeter-based and track-based isolation variables. . . . .  | 152 |
| 7.22 | The efficiency for electrons and muons applying different treatment of leptons within a jet. . . . .  | 154 |
| 7.23 | The fake rate for electrons and muons applying different treatment of leptons within a jet. . . . .   | 155 |
| 7.24 | The fake rate for electrons and muons applying different isolation criteria. .  | 157 |
| 8.1  | Lepton multiplicity and jet multiplicity. . . . .   | 161 |
| 8.2  | Lepton $p_T$ distributions for events with 3 or more isolated leptons. . . . .  | 162 |
| 8.3  | Leading jet $p_T$ after requiring 3 isolated leptons and Significance $S/\sqrt{B}$ for different minimum $p_T$ cuts on the leading jet. . . . .   | 163 |
| 8.4  | The $\cancel{E}_T$ and $M_{eff}$ distributions after requiring three or more isolated leptons and $p_T^{jet1} > 150$ GeV. . . . .   | 164 |
| 8.5  | $\cancel{E}_T$ after requiring 3 isolated leptons. Significance $S/\sqrt{B}$ for different $\cancel{E}_T$ cuts. . . . .   | 166 |
| 8.6  | $M_{l+l-}$ after requiring $N_l \geq 3$ and $M_{eff}$ after all event selection cut “3-leptons+ $\cancel{E}_T$ ” search. . . . .  | 166 |
| 8.7  | Significance $S/\sqrt{B}$ as function of integrated luminosity for “3-leptons+jet” and “3-leptons+ $\cancel{E}_T$ ” search. . . . .   | 168 |
| 8.8  | $M_{l+l-}$ and significance $S/\sqrt{B}$ as function of the mass window cut $ M_{l+l-} - M_Z $ . . . . .  | 168 |
| 8.9  | $\cancel{E}_T$ after requiring 3 or more isolated leptons and a $l^+l^-$ lepton pair with $ M_{l+l-} - M_Z  > 15$ GeV. . . . .  | 169 |
| 8.10 | $p_T$ of the leading jet after the $Z$ -mass window and $\cancel{E}_T$ cut. Significance $S/\sqrt{B}$ as function of the cut maximum cut on leading jet $p_T$ . . . . .   | 170 |
| 8.11 | $\text{IP}_N$ distribution after the $\cancel{E}_T$ cut (no jet veto). . . . .  | 170 |
| 8.12 | The discovery reach in terms of the statistical significance $S/\sqrt{B}$ as function of the integrated luminosity. . . . .   | 172 |
| 9.1  | $p_T$ for muons which are closer to a jet than $\Delta R = 0.4$ . . . . .   | 179 |
| 9.2  | $p_T^{jet}/p_T^\mu$ for muons which are closer to a jet than $\Delta R = 0.4$ . . . . .   | 179 |
| 9.3  | Comparison of the number of jets and jet $p_T$ in SU4 and $t\bar{t}$ . . . . .  | 181 |
| 9.4  | Comparison of $\Delta R$ between muons as the closest jet as reconstructed by the Anti- $K_T$ and fixed cone $\Delta R = 0.4$ algorithms. . . . .   | 181 |

|       |   |     |
|-------|---|-----|
| 9.5   | Comparison of the relative $p_T^{jet}/p_T^\mu$ for <i>prompt</i> muons and <i>extra</i> muons within a jet. . . . .                             | 182 |
| 9.6   | Comparison of the various measures of the statistical significance. . . . .   | 184 |
| 9.7   | Lepton $p_T$ distributions for events with 3 or more isolated leptons. MC simulation at 10 TeV. . . . .   | 185 |
| 9.8   | Trilepton events: $\cancel{E}_T$ and number of jets. . . . .  | 186 |
| 9.9   | Trilepton events: $p_T$ of the leading and sub-leading jet. . . . .   | 186 |
| 9.10  | Effective mass after jet $p_T$ cuts and $\cancel{E}_T$ cut. . . . .   | 188 |
| 10.1  | MSSM24 Directgaugino production, number of leptons and lepton $p_T$ . . . .   | 195 |
| 10.2  | MSSM24 Directgaugino production with decay via sleptons, number of leptons and lepton $p_T$ . . . . .   | 195 |
| 10.3  | Event selection variables directgaugino points. . . . .   | 197 |
| 10.4  | Event selection variables directgaugino points with decay via sleptons. . . .   | 198 |
| 10.5  | MSSM24 gauginoshake points, number of leptons and lepton $p_T$ number of jets. . . . .  | 202 |
| 10.6  | Event selection variables used jet inclusive search for gauginoshake points. .  | 203 |
| 10.7  | 2D significance plots for gauginoshake point 1 and 3. . . . .   | 205 |
| 11.1  | ATLAS event display of a jet event at 2.36 TeV. . . . .   | 208 |
| 11.2  | Comparison data-MC of jet $p_T$ and $\cancel{E}_T$ . . . . .  | 209 |
| 11.3  | Impact parameters $d_0$ , $z_0$ and $z_0 \sin \theta$ . . . . .   | 210 |
| 11.4  | Comparison data-MC: number of tracks per event, $p_T$ , $\eta$ , and $\phi$ . . . . .   | 211 |
| 11.5  | Comparison data-MC: number of tracks and track $p_T$ for events with $N_l \geq 1$ . .   | 211 |
| 11.6  | Comparison data-MC muon impact parameters. . . . .  | 212 |
| 11.7  | Comparison data-MC: number of pixel and SCT hits for muons. . . . .   | 212 |
| 11.8  | Comparison data - MC $p_T$ and $\eta$ distributions for combined and lowpt muons. .   | 213 |
| 11.9  | Comparison of data and MC calorimeter based isolation $E_{cal}^{\Delta R=0.X}$ for muons. .   | 215 |
| 11.10 | Comparison of data and MC track-based isolation $p_{T\Sigma track}^{\Delta R=0.X}$ for muons. . .   | 216 |
| 11.11 | Comparison of data and MC relative track-based isolation for muons. . . . .   | 216 |
| 11.12 | Comparison of data and MC track-based isolation for muons $p_{Tmax track}^{\Delta R=0.X}$ . .   | 217 |
| 11.13 | Comparison data-MC electron impact parameters. . . . .  | 218 |
| 11.14 | Comparison data-MC number of Pixel and SCT hits for electrons. . . . .  | 218 |
| 11.15 | Comparison of electron isEM tags loose, medium and tight. . . . .   | 219 |
| 11.16 | Comparison data-MC number of the $e/\gamma$ author tag. calorimeter-based isolation in cone $\Delta R = 0.3$ vs. $\eta$ . . . . .               | 219 |
| 11.17 | Comparison data-MC electron $p_T$ and $\eta$ distributions. . . . .   | 220 |
| 11.18 | Comparison of data and MC calorimeter-based isolation $E_{cal}^{\Delta R=0.X}$ for electrons from the <b>egamma</b> and soft algorithm. . . . . | 221 |
| 11.19 | Comparison of data and MC calorimeter based isolation $E_{cal}^{\Delta R=0.X}$ for electrons from the <b>egamma</b> and soft algorithm. . . . . | 222 |
| 11.20 | Comparison of data and MC track-based isolation $p_{T\Sigma track}^{\Delta R=0.X}$ for electrons. .   | 222 |
| 11.21 | Comparison of data and MC relative track-based isolation $p_{T\Sigma track}^{\Delta R=0.X}$ for electrons. . . . .                              | 223 |
| 11.22 | Comparison of data and MC track-based isolation $p_{Tmax track}^{\Delta R=0.X}$ for electrons. .  | 223 |
| 11.23 | $K_S^0$ mass and $E/p$ distribution for electron candidates matched to a track from a reconstructed $K_S^0$ . . . . .                           | 224 |



|       |   |     |
|-------|---|-----|
| 11.24 | calorimeter-based isolation $E_{\text{cal}}^{\Delta R=0.X}$ for electron candidates matched to a track from $K_S^0$ . . . . .   | 225 |
| 11.25 | Hadronic leakage in the first compartment of the hadronic calorimeter and lateral shower shape $R_\eta(37)$ for electron candidates matched to a track from $K_S^0$ . . . . . | 225 |
| 11.26 | Total shower width $w_{\text{tot1}}$ in the first compartment of the electromagnetic calorimeter. . . . .   | 226 |
| 12.1  | Graphical presentation of significances varying the three jet inclusive analysis cuts for SU4. . . . .  | 233 |
| 12.2  | SU4 $S_N$ and $Z_N$ plotted as function of the integrated luminosity. . . . .   | 242 |
| 13.1  | Cloud chamber used at several exhibitions and outreach booths. . . . .  | 253 |
| 13.2  | Mass break down handed out as a “weight certificate” at the Science Fair booth in September 2008. . . . .   | 255 |

# List of Tables

|     |   |     |
|-----|---|-----|
| 1.1 | Particles and their interactions . . . . .  | 9   |
| 3.1 | List of fields in the MSSM . . . . .  | 38  |
| 4.1 | ATLAS calorimeter system performance goals . . . . .  | 54  |
| 5.1 | NorduGrid resources 2006 - 2010 . . . . .   | 76  |
| 5.2 | Map of ARC job states to Eowyn states . . . . .   | 92  |
| 6.1 | Summary of parameters defining the ATLAS mSUGRA benchmark points. The list contains only points studied in this analysis. . . . .   | 123 |
| 6.2 | Particle mass spectrum (in GeV) for the SUSY benchmark points . . . . .   | 124 |
| 6.3 | List of the samples used in 14 TeV analysis, with cross sections $\sigma$ , NLO/LO $K$ -factors, average weights $\langle w \rangle$ related to the NLO, and corresponding integrated luminosities. . . . . | 128 |
| 7.1 | Fraction of muons passing the $p_T$ -dependent isolation cut. . . . .   | 141 |
| 7.2 | List of electron identification cuts applied to obtain a “medium” quality electron. . . . .   | 143 |
| 7.3 | Electron and muon efficiencies different treatments of leptons inside jets. . . . .   | 153 |
| 7.4 | Electron and muon fake rates different treatments of leptons inside jets. . . . .   | 154 |
| 7.5 | Electron and muon efficiencies for different isolation criteria. . . . .  | 156 |
| 7.6 | Electron and muon fake rates for different isolation criteria. . . . .  | 156 |
| 8.1 | Cut-flow table for trilepton jet inclusive analysis. . . . .  | 165 |
| 8.2 | Results jet inclusive analysis, $p_T^{\text{jet}} > 200$ GeV. . . . .   | 165 |
| 8.3 | Results jet inclusive analysis, $p_T^{\text{jet}} > 150$ GeV. . . . .   | 165 |

|      |   |     |
|------|---|-----|
| 8.4  | Results jet inclusive analysis, $\cancel{E}_T > 200$ GeV. . . . .   | 167 |
| 8.5  | Results jet inclusive analysis, $\cancel{E}_T > 150$ GeV. . . . .   | 167 |
| 8.6  | Results jet inclusive analysis, $\cancel{E}_T > 200$ GeV with $Z$ -mass window cut. . .   | 167 |
| 8.7  | Results jet inclusive analysis, $\cancel{E}_T > 150$ GeV with $Z$ -mass window cut. . .   | 167 |
| 8.8  | Cut-flow table jet exclusive trilepton analysis. . . . .  | 171 |
| 8.9  | Results jet exclusive analysis, SU2. . . . .  | 171 |
| 8.10 | Results jet exclusive analysis, SU2 - direct gaugino only. . . . .  | 171 |
| 9.1  | Comparison of cross sections for the dominant production processes for the SU3 benchmark point at 14 TeV and at 10 TeV. . . . .       | 176 |
| 9.2  | Comparison of cross sections for the dominant production processes for the SU4 benchmark point at 14 TeV and at 10 TeV. . . . .       | 176 |
| 9.3  | Summary of SM background samples at 10 TeV. . . . .   | 177 |
| 9.4  | Efficiencies and fake rates for combinations of lepton jet separation and additional kinematical cuts. . . . .                        | 180 |
| 9.5  | Efficiencies and fake rates for combinations of lepton jet separation and additional kinematical cuts using Anti- $K_T$ jets. . . . . | 183 |
| 9.6  | Cut flow for jet inclusive search at 10 TeV. Anti- $K_T$ tower jets. . . . .  | 189 |
| 9.7  | Cut flow jet inclusive search at 10 TeV. Cone Tower jets. . . . .   | 190 |
| 9.8  | Cut flow jet inclusive search at 10 TeV. Without K-factors. . . . .   | 190 |
| 9.9  | Cut flow jet inclusive search at 10 TeV. Relative track-based isolation. . .  | 190 |
| 9.10 | Cut flow jet inclusive search at 10 TeV. Tight calorimeter-based isolation. .   | 190 |
| 9.11 | Cut flow jet inclusive search at 10 TeV. SM multiplied by a factor of 2. . .  | 190 |
| 9.12 | Inregrated luminosity required for $5\sigma$ discovery jet inclusive analysis. . .  | 191 |
| 10.1 | Overview of the MSSM24 points with direct gaugino production. . . . .   | 194 |
| 10.2 | Significances for the jet exclusive search at 10 TeV, directgaugino points. .   | 197 |
| 10.3 | Significances for the jet exclusive search at 10 TeV, direct-gaugino-slepton points. . . . .  | 199 |
| 10.4 | Significances of the jet exclusive search at 10 TeV, directgaugino points. Track-based isolation. . . . .                             | 199 |
| 10.5 | Significances for the jet exclusive search at 10 TeV, direct-gaugino points with decay via sleptons. Track-based isolation. . . . .   | 200 |
| 10.6 | Overview of the MSSM24 gauginoshake points. . . . .   | 201 |
| 10.7 | Significances jet inclusive search at 10 TeV, gauginoshake points. . . . .  | 202 |
| 10.8 | Three dimensional cut scan 4-lepton gauginoshake signal and SM background. Cut ranges for background free signal. . . . .             | 204 |
| 10.9 | Significances for jet inclusive search at 10 TeV, gauginoshake points. . . .  | 204 |
| 11.1 | Breakdown of minimum bias event selection requirements for minimum bias 900 GeV data. . . . .   | 208 |
| 12.1 | Jet inclusive cut optimization for SU4, 10 TeV analysis. . . . .  | 231 |
| 12.2 | Variation of the jet inclusive analysis cuts for SU4. . . . .   | 232 |
| 12.3 | Summary of systematic effects for the SU4 jet inclusive analysis at 10 TeV.   | 236 |
| 12.4 | Summary of systematic effects for the SU3 jet inclusive analysis at 10 TeV.   | 236 |
| 12.5 | Summary of systematic effects for the SU1 jet inclusive analysis at 10 TeV.   | 237 |
| 12.6 | Variation of the jet inclusive event selection at 14 TeV. . . . .   | 238 |

|   |     |
|---|-----|
| 12.7 Overall summary of studied systematical effects for jet inclusive analysis<br>at 14 TeV. . . . . | 240 |
| 12.8 Variation of the jet exclusive event selection at 14 TeV. . . . .                                | 241 |
| 13.1 General evaluation of the Masterclass in 2005 and 2008. . . . .                                  | 253 |
| A.1 ATLAS external software . . . . .   | 261 |
| A.2 Proddb table ETASK . . . . .  | 262 |
| A.3 Proddb table EJOBDEF . . . . .  | 263 |
| A.4 Proddb table EJOBEXE . . . . .  | 263 |
| B.1 Branching ratios of sparticles in the SU1 benchmark point. . . . .                                | 265 |
| B.2 Branching ratios of sparticles in the SU2 benchmark point. . . . .                                | 266 |
| B.3 Branching ratios of sparticles in the SU3 benchmark point. . . . .                                | 266 |
| B.4 Branching ratios of sparticles in the SU4 benchmark point. . . . .                                | 266 |
| C.1 List of MC samples for 10 TeV analysis . . . . .  | 269 |
| C.2 QCD di-jet $p_T$ ranges corresponding to JX tags in MC sample names. . . .                        | 269 |
| C.3 List of MC MSSM24 samples for 10 TeV analysis . . . . .   | 270 |

